

THE
LONDON, EDINBURGH AND DUBLIN
PHILOSOPHICAL MAGAZINE
AND
JOURNAL OF SCIENCE

CONDUCTED BY

SIR OLIVER JOSEPH LODGE, D.Sc., LL.D., F.R.S.
SIR JOSEPH JOHN THOMSON, O.M., M.A., Sc.D., LL.D., F.R.S.
ALFRED W. PORTER, D.Sc., F.R.S.

AND

ALLAN FERGUSON, M.A., D.Sc.

“Nec araneorum sane textus ideo melior quia ex se fila gignunt, nec
noster vilior quia ex allenis libamus ut apes.”

JUST. LIPS. *Polit. lib. i. cap. l. Not.*

VOL. XXVII.—SEVENTH SERIES

JANUARY—JUNE 1939

LONDON:

TAYLOR & FRANCIS, LTD., RED LION COURT, FLEET STREET, E.C. 4.

Agents: WM. DAWSON & SONS, LTD., CANNON HOUSE, PILGRIM STREET, E.C. 4.

“Meditationis est perscrutari occulta; contemplationis est admirari perspicua . . . Admiratio generat quæstionem, quæstio investigationem, investigatio inventionem.”—*Hugo de S. Victore.*

— “Cur spirent venti, cur terra dehiscat,
Cur mare turgescat, pelago cur tantus amaror,
Cur caput obscura Phœbus ferrugine condât,
Quid toties diros cogat flagrare cometas,
Quid pariat nubes, veniant cur fulmina cœlo,
Quo micet igne Iris, superos quis conciat orbes
Tam vario motu.”

J. B. Pinelli ad Mazonium.



CONTENTS OF VOL. XXVII. (Nos. 180-185)

(SEVENTH SERIES)

NUMBER 180—JANUARY

	Page
I. Electronic Waves. By Sir J. J. THOMSON, O.M., F.R.S. (Plate I)..	1
II. The Concept and Determination of Mass in Newtonian Mechanics. By V. V. NARLIKAR, Benares Hindu University, India	33
III. The Flexure of an Infinite Elastic Strip on an Elastic Foundation. By GEOFFREY BOSSON, M.Sc.	37
IV. A further Note on the Definition and Determination of Mass in Newtonian Mechanics. By C. G. PENDSE, M.A., Ph.D.	51
V. Three Elementary Examples of the Uncertainty Principle. By D. S. KOTHARI, Ph.D. (Cambridge)	62
VI. A Hot Wire Method for the Thermal Conductivities of Gases. By G. G. SHERRATT, B.A., and EZER GRIFFITHS, D.Sc., F.R.S., Physics Department, National Physical Laboratory, Teddington, Middlesex	68
VII. On the Neutron-Proton Scattering. By Prof. K. C. KAR, D.Sc., and D. BASU, B.Sc.	76
VIII. The Elementary Particle. By Professor W. WILSON, F.R.S., and Miss J. CATTERMOLÉ, Ph.D.	84
IX. A Note on the Method of Parallax and the Resolving Power of the Eye. By SUKHDEO BIHARI MATHUR, M.Sc., Hindu College, University of Delhi, Delhi	94
X. The Geometry of Discrete Vector Maps. By D. M. WRINCH, M.A., D.Sc., Mathematical Institute, Oxford	98
XI. Symmetrical Figures on Circular Plates and Membranes. By R. C. COLWELL, Ph.D., J. K. STEWART, Ph.D., and A. W. FRIEND, M.S., West Virginia University. (Plate II.)	123
XII. Notices respecting New Books :—	
M. M. FRÉCHET's <i>Traité du Calcul des Probabilités et de ses Applications. Tome i. fascicule 3: Recherches Théoriques Modernes</i>	129
Dr. W. EDWARDS DEMING's <i>Some Notes on Least Squares</i>	130
Mr. A. J. ANSLEY's <i>An Introduction to Laboratory Technique</i> ...	130
Mr. C. N. MOORE's <i>Summable Series and Convergence Factors</i> ...	131
Mr. DONALD G. FINK's <i>Engineering Electronics</i>	131
Dr. P. WESSEL's <i>Physik für Studieren an Technischen Hochschulen und Universitäten</i>	132
Mr. E. HIRSCHLAFF's <i>Fluorescence and Phosphorescence</i>	132

NUMBER 181—FEBRUARY

	Page
XIII. The Force-time Law governing the Impact of a Hammer on a Stretched String. By N. DAVY, J. H. LITTLEWOOD, and M. McCAIG, University College, Nottingham. (Plate III.)	133
XIV. The Variation of Solar Ultra-Violet Radiation during the Sunspot Cycle. By E. V. APPLETON, F.R.S., and R. NAISMITH, A.M.I.E.E.	144
XV. Euler's Problem of Two Fixed Centres of Gravitation. By MONROE H. MARTIN	149
XVI. On the Critical Dimensions of Tuned Transmitting Circular Loop Aerials. By S. S. BANERJEE, D.Sc., Department of Physics, Benares Hindu University	174
XVII. On the Problem of Wave-motion for Sub-Infinite Domains. By Dr. ARNOLD N. LOWAN	182
XVIII. Studies of the Tone Quality of Organ-Pipes.—II. Reed-Pipes. By M. MOKHTAR, M.Sc. (King's College, Newcastle-on-Tyne). (Plates IV. & V.)	195
XIX. Hydrodynamic Forces on an Accelerated Cylinder moving in Two-dimensions. By M. A. OMARA, Docteur-ès-Sciences Mathématiques, Lecturer in Applied Mathematics, The Egyptian University, Cairo (Egypt).	200
XX. Uniform Motion of a Sphere or a Cylinder through a Viscous Liquid. By B. R. SETH, M.A., D.Sc.	212
XXI. A Geometric Derivation of the Second Order Wave Equation. By H. W. HASKEY, Ph.D.	221
XXII. The Lift and Drag of a Rectangular Wing spanning of a Free Jet of Circular Section. By H. B. SQUIRE, Manchester University	229
XXIII. On the Development of Turbulent Liquid Motion over an Infinite Plate. By MANOHAR RAY, Lahore, India	240
XXIV. Geiger Point Counters. By B. DASANNACHARYA, M.A., Dr.Phil., F.Inst.P., Professor of Physics, Benares Hindu University, Benares, India, and AMAR CHAND SETH, M.Sc. (Benares)	249
XXV. Notices respecting New Books:—	
Mr. F. E. HOARE's A Textbook of Thermodynamics	258
Prof. W. M. SMART's Stellar Dynamics	258
Prof. E. SUTTON's Demonstration Experiments in Physics	260

NUMBER 182—MARCH

XXVI. Further Note on the Propagation of Radio Waves over a Finitely Conducting Spherical Earth. By BALTH. VAN DER POL, D.Sc., and H. BREMMER, D.Sc., Natuurkundig Laboratorium der N. V. Philips' Gloeilampenfabrieken, Eindhoven, Holland	261
XXVII. The Thermal Stability of a Cylindrical Stratified Dielectric. By S. WHITEHEAD, M.A., Ph.D.	276
XXVIII. Some Problems of Finite Strain.—I. By B. R. SETH, M.A., D.Sc.	286
XXIX. Factors affecting the Limit of Solubility of Elements in Copper and Silver. By Professor E. A. OWEN, M.A., Sc.D., and the late EDGAR WYNNE ROBERTS, M.Sc., University College of North Wales, Bangor	294
XXX. The Theorem of Three Moments with Variable "I" but without Thrust. By F. J. TURTON, B.Sc., A.R.C.Sc., City and Guilds College	328

	Page
XXXI. Theory of the Fractionation of Gaseous Mixtures by Diffusion. The Characteristics of the Hertz-Mercury-Vapour-Pump. By E. BLUMEN-THAL.....	341
XXXII. Entropy of Fermi-Dirac Gas. By F. C. AULUCK, Panjab University, Lahore	370
XXXIII. Hamilton's Quaternions and Minkowski's Potentials. By OTTO F. FISCHER.....	375
XXXIV. Notices respecting New Books :—	
Mr. V. M. FAIRES's Applied Thermodynamics	386
Mr. V. M. FAIRES's Elementary Thermodynamics	386
Messrs. D. JORDAN LLOYD and D. SHORE's Chemistry of Proteins	387
Obituary :—Prof. A. W. PORTER, D.Sc., F.R.S.	388

NUMBER 183—APRIL

XXXV. The Calculation of Equilibrium Internuclear Distance for Diatomic Hydrogen, Hydrides, and Deuterides in Ground and Excited States. By C. H. DOUGLAS CLARK, D.Sc., and JOHN L. STOVES, Ph.D....	389
XXXVI. An Auxiliary Equation for use with the Heaviside Expansion Theorem. By W. B. COULTHARD, B.Sc.(Eng.) London, A.M.I.E.E.	404
XXXVII. Approximation Formulæ for a well-known Difference of Products of Two Cylinder Functions. By H. BUCHHOLZ	407
XXXVIII. Diffraction and Refraction of a Horizontally Polarized Electromagnetic Wave over a Spherical Earth. By MARION C. GRAY....	421
XXXIX. Direct Determination of Stresses from the Stress Equations in some Two-Dimensional Problems of Elasticity.—Part II. Thermal Stresses. By BIBHUTIBHUSAN SEN, Krishnagar College, Bengal	437
XL. A Graphical Construction for Stress. By JAMES J. GUEST	445
XLI. Some Problems of Finite Strain.—II. By B. R. SETH, M.A., D.Sc.	449
XLII. Quantum Theory of Gravitation. By A. KALIVARIAS, D.Sc., O. A. Stathatos Research Fellow	453
XLIII. Non-linear Theory of Elasticity and the Linearized Case for a Body under Initial Stress. By Prof. M. A. BIOT.....	468
XLIV. Vector Maps of Finite and Periodic Point Sets. By D. M. WRINCH, M.A., D.Sc., Mathematical Institute, Oxford.....	490
XLV. The Stress Distribution in an Aeolotropic Circular Disk. By H. ÔKUBO	508
XLVI. Notices respecting New Books :—	
Dr. LUDWIG BERGMANN's Ultrasonics and their Scientific and Technical Applications	513
Prof. RICHARD C. TOLMAN's The Principles of Statistical Mechanics	514
Prof. C. E. VAN HORN's A Preface to Mathematics	515
Mr. J. M. THOMAS's Theory of Equations	515
Prof. A. ADRIAN ALBERT's Modern Higher Algebra	515
Prof. FARRINGTON DANIELS's Lectures on Chemical Kinetics ..	516

NUMBER 184—MAY

	Page
XLVII. The Diffraction of Wireless Waves round the Earth. (A Summary of the Diffraction Analysis, with a Comparison between the Various Methods.) By G. MILLINGTON, Research and Development Department, Marconi's Wireless Telegraph Co., Ltd., Chelmsford, Essex	517
XLVIII. A Kinematical Description of a Flat Space-Time. By C. GILBERT, Ph.D., King's College, Newcastle	543
XLIX. The Atomic Heat of Potassium. By L. G. CARPENTER and C. J. STEWARD, University College, Southampton	551
L. Optical Dispersion and the Vibratory Doublet Photon. By E. TAYLOR JONES, D.Sc., Professor of Natural Philosophy in the University of Glasgow	565
LI. On some Geophysical Consolidation Problems. By H. Löwy, Dr.Phil.	576
LII. On Dixon's Formula for well-poised Series. By T. M. MACROBERT, Professor of Mathematics in the University of Glasgow	579
LIII. Thermal Instability of Dielectrics for alternating Voltages when the Loss Angle is dependent upon the Field Strength. By A. GEMANT, University of Wisconsin, U.S.A., and S. WHITEHEAD	582
LIV. Direct Determination of Stresses from the Stress Equations in some Two-Dimensional Problems of Elasticity.—Part III. Problems of Non-Isotropic Material. By BIBHUTIBHUSAN SEN, Krishnagar College, Bengal	596
LV. A Matrix Theory Development of the Theory of Symmetric Components. By N. S. RISLEY, M.S., and Professor RICHARD S. BURINGTON, Ph.D., M.A., Department of Mathematics, Case School of Applied Science, Cleveland, Ohio	605
LVI. The Equilibrium Diagram of Iron-nickel Alloys. By Prof. E. A. OWEN, M.A., Sc.D., and A. H. SULLY, M.Sc., University College of North Wales, Bangor. (Plates VI. & VII.)	614
LVII. The Disintegration of Erythrocytes and Denaturation of Hemoglobin by High Pressure. By R. B. DOW, A.M., Ph.D., and J. E. MATTHEWS, Jr., A.B., Departments of Physics and Agricultural and Biological Chemistry, The Pennsylvania State College. (Plates VIII.—X.)	637
LVIII. Equivalence of two Piezoelectric Oscillating Quartz Crystals of Symmetrical Outlines with Respect to a Plane Perpendicular to an Electrical Axis. By ISSAC KOGA, Doctor of Engineering (Tokyo University of Engineering)	640
LIX. Notices respecting New Books :—	
Prof. ALEXANDER FINDLAY'S The Phase Rule and its Applications.	644
Mr. J. C. SPEAKMAN'S Modern Atomic Theory	644

NUMBER 185—JUNE

	Page
LX. On the General Validity of Nyquist's Theorem. By D. A. BELL, B.A., B.Sc.	645
LXI. The Dielectric Constant of Benzene. By WILFRED C. VAUGHAN, B.Sc., Ph.D.	661
LXII. The Dielectric Constant, Dipole Moment, and Molecular Polarization of 1.4 Dioxane ($C_4H_8O_2$). By WILFRED C. VAUGHAN, B.Sc., Ph.D.	669
LXIII. The Thermal and Electrical Conductivities of some Magnesium Alloys. By R. W. POWELL, B.Sc., Ph.D., Physics Department, The National Physical Laboratory, Teddington, Middlesex	677
LXIV. A Note on the Nature of Sliding Friction. By J. J. BIKERMAN and E. K. RIDEAL, Department of Colloid Science, The University, Cambridge	687
LXV. The Interpretation of the Michelson-Morley and Kennedy-Thorndike Experiments. By HERBERT DINGLE, A.R.C.S., D.Sc., Professor of Natural Philosophy, Imperial College of Science and Technology	693
LXVI. Some Formulæ for the Associated Legendre Functions of the First Kind. By T. M. MACROBERT, Professor of Mathematics, University of Glasgow	703
LXVII. Note on a Problem in Heat Conduction. By HAROLD W. WOOLLEY	706
LXVIII. Calculation of Triode Constants. By J. H. FREMLIN, M.A., Ph.D., A.Inst.P.	709
LXIX. The Evidence for a Superlattice in the Nickel-iron Alloy Ni_3Fe . By P. LEECH, B.Sc., and C. SYKES, D.Sc.	742
LXX. Thermodynamical Properties of some Supraconductors. By J. G. DAUNT, A. HORSEMAN, and K. MENDELSSOHN, Clarendon Laboratory, Oxford	754
LXXI. The General Solutions to Problems in Heat Transmission between two Media, one at least of which is a Fluid, the Temperature Gradient for one of the Media along the Line of Flow being known. By W. J. WALKER, D.Sc., Ph.D., University of the Witwatersrand, Johannesburg, South Africa	765
LXXII. Errata to Paper "On the Problem of Wave-motion for Sub-Infinite Domains."—February 1939, p. 182. By ARNOLD N. LOWAN	769
LXXIII. Notices respecting New Books :—	
Messrs. G. H. HARDY and E. M. WRIGHT's An Introduction to the Theory of Numbers	770
Mr. EDWIN HERBERT HALL's A Dual Theory of Conduction in Metals	771
Mr. D. SHOENBERG's Superconductivity	772
Index	773

PLATES

- I. Illustrative of Sir J. J. THOMSON's Paper on Electronic Waves.
 - II. Illustrative of Drs. R. C. COLWELL and J. K. STEWART and Mr. A. W. FRIEND's Paper on Symmetrical Figures on Circular Plates and Membranes.
 - III. Illustrative of Messrs. N. DAVY, J. H. LITTLEWOOD, and M. McCaig's Paper on the Force-time Law governing the Impact of a Hammer on a Stretched String.
 - IV. & V. Illustrative of Mr. M. MOKHTAR's Paper on Studies of the Tone Quality of Organ-Pipes.—II. Reed-Pipes.
 - VI. & VII. Illustrative of Prof. E. A. OWEN and Mr. A. H. SULLY's Paper on the Equilibrium Diagram of Iron-nickel Alloys.
 - VIII.—X. Illustrative of Dr. R. DOW and Mr. J. E. MATTHEWS's Paper on the Disintegration of Erythrocytes and Denaturation of Hemoglobin by High Pressure.
-

THE
LONDON, EDINBURGH & DUBLIN
PHILOSOPHICAL MAGAZINE
AND
JOURNAL OF SCIENCE

[SEVENTH SERIES—VOL. 27]

I. *Electronic Waves.*

By Sir J. J. THOMSON, O.M., F.R.S.*

[Received October 21, 1938.]

[Plate I.]

THAT electrons may be and, as far as we know, always are accompanied by trains of waves has been proved conclusively by experiments on electronic diffraction. Very little, however, is known about the nature of these waves; are they electromagnetic waves governed by Maxwell's equations, or are they something *sui generis*?

A fundamental difference between waves of light and electronic waves was discovered by my son, Professor G. P. Thomson, at the very beginning of his experiments on these waves †. He observed that when electrons passed through very thin sheets of metal diffraction patterns were produced similar to those which might have been produced if waves of light of appropriate wave-length had passed through the plate. By measuring the diffraction patterns produced by the electrons he determined the wave-length of light that would have produced them, and

* Communicated by the Author.

† Proc. Roy. Soc. London, A, pp. 119-651.

found that the product of this wave-length and the velocity of the electron was a constant quantity. But whereas the diffraction pattern due to light would not be affected by magnetic forces, that due to the electrons was displaced by these forces, and the deflexion was in the same direction as that of an electron travelling with the electronic waves. This occurred when the deflecting magnet was far away from the metal foil where the diffraction was produced, showing I think conclusively that the electronic waves must be accompanied by an electron. This, however, raises a very interesting difficulty; the velocity of the electron is generally much less than the velocity of light which we take to be that of propagation of the waves, so that if the waves are to keep in touch with the electrons they must pursue a very circuitous path.

Again, cathode rays can travel long distances through space devoid of matter without losing energy. Thus all the energy must travel in the direction in which the waves are travelling, and there can be no radiation at right angles to this direction. In this respect the electronic system resembles a quantum of light which can also travel through space without losing energy. Several types of trains of waves possess this property; one of these is that of a train of "cylindrical waves." Such a train consists of certain distributions of electric and magnetic forces round a central axis parallel to the direction in which the waves are travelling.

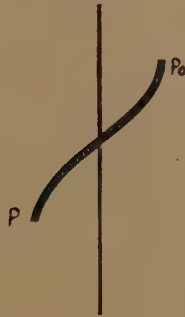
We shall have to consider these later on in more detail: now we shall confine ourselves to the description of a very simple case. In this the lines of electric force are circles with their centres along the axis of the train and their planes at right angles to it. The lines of magnetic force are in planes passing through the axis of the train and are symmetrical about this axis. From the symmetry of Maxwell's equation with respect to electric and magnetic forces if we have got one solution we can get another by interchanging the electric and magnetic forces, so that we might have circular lines of magnetic force and lines of electric force in the meridians. From such trains of waves there is no loss of energy by radiation even though the length of the train is a large multiple of the length of the waves in the train; the sharpness of the photographs of the diffraction patterns show that this must be the case for the electronic waves. The energy in cylindrical waves diminishes but slowly as the distance from the axis increases, so that they would have a wave front of considerable area. This is also demanded by the sharpness of the photographs.

We come back to the difficulty of the waves keeping in touch with the slower electrons. The simplest solution which has occurred to me is to suppose that the axis of the *waves*, instead of being a straight line as in the case just considered, is one which twists round the straight line which is the path of the electron (fig. 1). It will be shown later on that if the

axis of the curve twisting round the path of the electron is a regular helix there will be no loss of energy by radiation. On this view the electron and the waves which accompany it are rotating about the line along which the electron is moving, and the whole system, electron and waves, is moving like a screw whose axis is the path of the electron. The velocity of the screw along its axis and the angular velocity around it are the same as that of the electron, while the velocity of the electronic waves along their spiral path is equal to the velocity of light. This conception seems to be in accordance with that of the "spinning electron" provided that by "electron" we understand not merely the electric charge but a system which includes the waves as well.

Let s be the distance P_0P measured along the arc of the spiral of a point P from a fixed point P_0 on the spiral passing through P ; then the expressions for the electric and magnetic forces in the waves will all

Fig. 1.



contain the factor $e^{i(pt - qs)}$, where $2\pi/p$ is the frequency of the waves and p/q their velocity, which we take to be c , that of light. If z, ρ, θ are cylindrical coordinates of the point P , the axis of z being the line along which the electron is moving, then if α be the angle which ds an element of the spiral makes with the axis of z ,

$$s = z \cos \alpha + \rho \theta \sin \alpha, \\ e^{i(pt - qs)} = e^{i(pt - q(z \cos \alpha + \rho \theta \sin \alpha))}. \quad \dots \dots (1)$$

Since the electron keeps in touch with the waves the velocity of the waves parallel to z must be equal to the velocity of the electron.

Since the velocity of the waves along the arc of the spiral is c , if u be the velocity of the electron, $u = c \cos \alpha$. From equation (1) the wavelength λ measured in the direction of motion of the electron is given by the equation

$$\lambda = 2\pi/q \cos \alpha = 2\pi c/uq,$$

or since

$$q = p/c,$$

$$u\lambda = 2\pi c^2/p.$$

Since the experiments made on electronic waves show that $u\lambda$ is constant, this equation shows that p is constant and hence that all electronic waves have the same frequency. This fits in well with the conception of the spinning electron, and suggests that this frequency is the frequency of the spin. The experiments showed that $u\lambda$ was approximately 7.5; this makes $2\pi/p$, the frequency of the spin, equal to 1.2×10^{20} . The equation $h\nu = mc^2$, the well known equation of relativity, would give $hp/2\pi = mc^2$, where mc^2 is the mass of the electronic system. This is the energy of the system consisting of the spinning electron together with that of the waves associated with it. We can form an estimate of the energy in the waves of the electronic system by finding what the frequency of a wave must be for a quantum of the wave to have energy equal to that possessed by the whole system. From the preceding equations we see that the frequency of the wave must be 1.2×10^{20} ; this corresponds to a wave-length 2.5×10^{-10} . This is smaller than all known wave-lengths except those of the very hardest γ -rays. It is much smaller than the wave-lengths observed by Davisson and Germer in the first experiments made on electronic waves, which ranged from $.5 \times 10^{-8}$ to 1.5×10^{-8} , when the energy of the cathode rays ranged from 600 to 65 electron volts. The energy in a quantum of the electronic waves will therefore be less than the energy in the electronic system, and there must then be a part, and possibly the greater part, of the energy in the electron itself due to its spin and to its charge of electricity.

Size of the Electron.

The electronic system on this view will have for its outer boundary the cylinder round which the helical electronic waves are wound, and we may take the radius of this cylinder as a measure of the radius of the electron. If α is the angle the tangent to the helix makes with the axis of the cylinder on which it is wound, a the radius of the cylinder, ds an element of the spiral, and $d\phi$ the angle between the radii perpendicular to the axis at the ends of this element, then

$$ds \sin \alpha = a d\phi;$$

thus s is proportional to ϕ , and when ϕ is equal to 2π must be equal to the wave-length, so that

$$\lambda \sin \alpha = 2\pi a,$$

but

$$c \cos \alpha = u,$$

so that

$$2\pi a = \lambda \left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}};$$

since

$$\lambda u = 2\pi c^2/p, \text{ and } 2\pi/p = .8 \times 10^{-20},$$

$$2\pi a = 2.4 \frac{c}{u} \left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}} \times 10^{-10}.$$

The values of $2\pi a$ corresponding to various values of u are given in the following table; V is the potential difference in volts required to impart the velocity u to the electron.

$u.$	$2\pi a.$	$V.$
$c/10^3$	2.4×10^{-7}	.25
$c/10^2$	2.4×10^{-8}	25
$c/10$	2.4×10^{-9}	2500
$c/5$	2.4×10^{-10}	100,000

We see from this table that the effective size of electrons is greater when the velocity is small than when it is great, and that unless V is greater than .25 volt the electronic system would be larger than an atom, so that we should not expect it to possess the high penetrating power which is possessed by electrons with higher velocities. We see too that when the velocity of the electron approaches the velocity of light the size of the electron diminishes to a small fraction of that of an atom and approaches the values that have usually been assigned to it.

The method we have used in obtaining these equations is entirely kinematical; we have not used any equations of the electromagnetic field, but merely assumed that the velocity of the waves in the direction in which the electron is moving is the same as the velocity of the electron. This is exactly what is done in the theory of relativity when a vibration which when referred to fixed axes is represented by $\cos pt$ is, when referred to axis moving with a velocity u , represented by $\cos p\left(t - \frac{ux}{c^2}\right)$. The wave-length λ of this is given by the equation

$$\frac{2\pi}{\lambda} = \frac{up}{c^2} \quad \text{or} \quad u\lambda = \frac{2\pi c^2}{p},$$

which is the same as that given on p. 3.

There are other frequencies connected with an electron besides that of its spin. Thus, if we regard the electron as an electrified sphere the lines of electric force starting from it will not be in equilibrium unless they are uniformly distributed over the surface of the sphere. If this distribution be disturbed by an electric charge passing past the electron it will not be in equilibrium after this charge has moved away, but will oscillate about the uniform distribution. I have discussed the theory

of these oscillations *, and shown that if the vibrations are represented by e^{ipt} the values of pa/c , a being the radius of the sphere and c the velocity of light, are given by transcendental equations whose roots were not wholly real but contain an imaginary part. This implies a very rapid decay of the vibrations with the time, the vibrations in many cases dying to a small fraction of their original amplitude before the first vibration was completed. As the energy is dissipated by radiation in a small fraction of the time of vibration of the waves, the energy is dissipated as a pulse and not as a train of waves.

Cylindrical Waves.

Before discussing the problem of helical waves, which we have supposed to represent electronic waves, it will be helpful to consider the special case when the velocity of the electron is so nearly that of light that the waves are travelling in straight lines parallel to the path of the electron. It is appropriate to use for this investigation cylindrical coordinates ρ, θ, z with this line for the axis of z .

Maxwell's Equations expressed in Terms of Cylindrical Coordinates.

If P, Q, R are the components of electric force along ρ, θ, z ; α, β, γ those of the magnetic force, K and μ respectively the specific inductive capacity and magnetic permeability of the medium through which the waves are passing, then

$$\begin{aligned} K \frac{dR}{dt} &= \frac{1}{\rho} \frac{d}{d\rho} (\rho\beta) - \frac{1}{\rho} \frac{d\alpha}{d\theta}, \\ K \frac{dQ}{dt} &= \frac{d\alpha}{d\beta} - \frac{d\gamma}{d\rho}, \\ K \frac{dP}{dt} &= \frac{1}{\rho} \frac{d\gamma}{d\theta} - \frac{d\beta}{dz}, \end{aligned}$$

together with a similar set of equations where P, Q, R and α, β, γ are interchanged as are also K and μ . If the waves are travelling in the direction of z and are proportional to $e^{i(pt-qz)}$ we find after some reductions

$$\frac{1}{c^2} \frac{d^2 R}{dt^2} = \frac{d^2 R}{dz^2} + \frac{1}{\rho} \frac{d}{d\rho} \left(\rho \frac{dR}{d\rho} \right) - \frac{1}{\rho^2} \frac{d^2 R}{d\theta^2},$$

where $c^2 = 1/K\mu$.

If R varies as $e^{m\theta}$, this becomes

$$\left(\frac{p^2}{c^2} - m^2 \right) R + \frac{1}{\rho} \frac{d}{d\rho} \left(\rho \frac{dR}{d\rho} \right) - \frac{n^2 R}{\rho^2} = 0;$$

putting $\rho \sqrt{p^2/c^2 - m^2} = x$, this becomes

$$R + \frac{1}{x} \frac{d}{dx} \left(x \frac{dR}{dx} \right) - \frac{n^2}{x^2} R = 0.$$

* 'Recent Researches in Electricity and Magnetism,' p. 364.

The solution of which when x is real is

$$R = e^{i(pt - mz + n\theta)} \times J_n(x), \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where $J_n(x)$ is Bessel's Function of order n . Since

$$m = q \cos \alpha = qu/c = pu/c^2,$$

where u is the velocity of the electron.

$$x = \rho \sqrt{p^2/c^2 - q^2 a^2},$$

and since $q = p/c$

$$x = (\rho p/c) \sqrt{1 - u^2/c^2}.$$

The "phase velocity" of the wave in the direction in which it is travelling is p/m , so that unless this is greater than c , x is not real and the solution does not apply.

Equation (1) was obtained by Dr. Japolski*, who has made very extensive and interesting investigations on cylindrical waves. He gets over the difficulty by supposing that the "phase velocity" can be greater than that of light, and that it is so in the case of his cylindrical waves. The difficulty does not arise in the case we have been considering, for the wave is not travelling in the direction of z but in one inclined to it, and its phase velocity along z is only a component of its phase velocity c in the direction in which the wave itself is travelling. Thus the phase velocity in the direction in which the electron is travelling is always greater than the velocity of light.

The expressions of the electric and magnetic forces in these electronic waves are of the form

$$\phi = \chi J_n(x),$$

where

$$\chi = e^{i(p\theta - mz)}.$$

The Bessel functions do not diminish rapidly as x increases; for large values of x they are proportional to $1/\sqrt{x}$. The energy in unit volume due to the electric and magnetic forces is proportional to the square of these forces, i. e., to $1/x$. If we draw a series of annuli bounded by circles whose radii are in arithmetical progression the area of each annulus will be proportional to x , and since the energy is approximately proportional to $1/x$ the amount of energy passing through each annulus will only diminish slowly; thus the dimensions of the wave train at right angles to its direction of propagation will be considerable, which will make for good definition in the photographs of the diffraction pattern.

No Loss of Energy by Radiation.

When a wave is travelling parallel to the axis of z —the direction of motion of the electron,—the expressions for the electric and magnetic

* Japolski, Phil. Mag. p. 934 (May 1935).

forces will contain a factor of the form $e^{i(2t-mz)} \equiv \epsilon^{i\psi}$. Some of these quantities will contain the factor $\cos \psi$, others $\sin \psi$. If one of the terms in Maxwell's equations contains $e^{i\psi}$ in the form $\cos \psi$ the others must do so too. Suppose that R is proportional to $\cos \psi$, then the differential coefficients of R with respect to t or z will be proportional to $\sin \psi$, those with respect to the other coordinates to $\cos \psi$. If R is proportional to $\cos \psi$ we see from the equation for dR/dt that α and β must be proportional to $\sin \psi$. From the equation for dQ/dt we see that Q and α must have the same phase $\sin \psi$, and γ the phase $\cos \psi$, and from the equation for dP/dt we see that P must have the same phase as β . Thus γ and R are proportional to $\cos \psi$ and Q, P, α, β to $\sin \psi$. The flow of energy is by Poynting's Theorem the vector product of the electric and magnetic forces, hence the radial flow of energy, *i. e.*, that along ρ is equal to $(R\beta - Q\gamma)$, this will contain the term $\sin \psi \cos \psi$, *i. e.*, $\cos (pt - mz) \sin (pt - mz)$ the mean value of this taken with respect to either t or z is zero, and so therefore is the flow of energy along the direction at right angles to that in which the wave is travelling. Thus the system will not lose energy as it passes through space and the number of wave-lengths in the train will not diminish. The flow of energy in the direction of propagation of the waves is proportional to $Q\alpha - P\beta$ or to $\sin^2 \psi$; the average value of this is finite.

Since there is no loss of energy by radiation the train of waves might be one with a considerable number of wave-lengths, which is in accordance with the sharpness of the definition in the photographs of the diffraction patterns due to electronic waves.

Two Types of Light Quanta and of Electronic Waves.

We see from Maxwell's equations that when $n=0$, *i. e.*, when the waves are symmetrical about the axis,

$$\gamma = A \sin (pt - mz) J_0(x),$$

where

$$x = \frac{2\pi\rho}{\lambda} \sqrt{1 - \frac{u^2}{c^2}},$$

would be a solution of the equations; putting this value of γ in Maxwell's equations we get

$$Q = \frac{\mu p}{mv} A \cos (pt - mz) J_1(x),$$

$$\alpha = \frac{\mu}{v} A \cos (pt - mz) J_1(x);$$

we cannot, however, determine the values of P, R , or β in terms of A .

If we put

$$R = B \sin (pt - mz) J_0(x),$$

we find

$$P = \frac{B}{v} \cos (pt - mz) J_1(x),$$

$$\beta = \frac{Kp}{vm} B \cos (pt - mz) J_1(x),$$

but we cannot find the values of Q , α , γ in terms of B . The six quantities P , Q , R , α , β , γ form two sets, one being Q , α , γ , the other P , R , β . If we know the value of one of the quantities in a set we can find the other two in that set, but we cannot find any in the other set. This applies assuming the electromagnetic theory of light to light quanta as well as electronic waves. So that on this view there might be two kinds of light of the same wave-length.

It is to be noticed that in each type there is a longitudinal component of the wave; in type A it is the magnetic component, in type B the electric.

The absence of the longitudinal component is not a fundamental condition for wave propagation, it only applies to the case of plane waves. It is on quite a different footing from the one that in space devoid of electric charges or magnetic poles the lines of electric and magnetic force must form closed curves. In the first type of quantum on page 8 the lines of electric force form closed curves, but the lines of magnetic force could not do so unless they had a component in the direction of propagation.

We can see too in another way that there must be a longitudinal component in case 1, for the line integral of Q round its circle is finite, and this must equal the rate of change of the number of lines of magnetic force passing through the circle. This, however, could not be finite unless there were magnetic forces perpendicular to the plane of this circle.

There is no evidence from optics of the existence of two different quanta of light of the same wave-length. On the quantum theory of light, however, we should not expect to be able to discover such a difference by optical experiments any more than that we should be able to discover that air contained molecules of different kinds by experiments on the reflexion and refraction of sound. The equations for waves of light as well as those of sound are statistical equations between average values, and do not tell us anything about the behaviour of individuals.

To take an extreme case, even if every quantum of type (1) that struck a reflecting surface passed through it there would still be reflexion from the surface. For suppose the incidence is perpendicular, and that in the incident beam there are n quanta per unit volume, then the number of quanta which passed across unit area of the reflecting surface per second would be cn , c being the velocity of light. After they cross the reflecting surface, the velocity of the quanta is c/μ , where μ is the refractive index of the medium; so that the stream passing through unit area per sec.

would now only be nc/μ ; this is less than the stream flowing in, there would therefore be a continuous increase in the number of quanta in the refracting medium. To relieve this some of the quanta diffuse backwards, recross the surface of separation, and produce a reflected stream from the surface equal to $cn - cn/\mu$, the stream in the refracting medium being cn/μ . Thus the existence of both reflexion and refraction does not necessarily imply that a quantum has been split into two parts, one of which is reflected the other refracted.

Helical Waves.

For the discussion of these it is convenient to use curvilinear coordinates, ζ, η, ξ ; the axis of ζ is the tangent to the helical axis of the waves, the axis of η is the principal normal of the helix; this at any point coincides with the radius drawn from that point to the straight axis of the helix; the axis of ξ is the binormal of the helix. If ν, μ, λ are the direction cosines of these axes referred to axes z, y, x , when the axis of z is the line along which the electron is moving, and y and x are axes at right angles to each other and to the axis of z , then if $y = r \sin \theta, x = r \cos \theta$, and α the constant angle the tangent to the helix make with the axis of z , we get

for the axis of ζ : $\nu = \cos \alpha$; $\mu = -\sin \alpha \cos \theta$; $\lambda = \sin \alpha \sin \theta$;

for the axis of η : $\nu = 0$; $\mu = \sin \theta$; $\lambda = \cos \theta$;

for the axis of ξ : $\nu = \sin \alpha$; $\mu = \cos \alpha \cos \theta$; $\lambda = -\cos \alpha \sin \theta$.

If $\delta x, \delta y, \delta z$ are the differences of the coordinates of two points close together

$$\delta x = \cos \alpha \delta \rho + \sin \alpha \delta \xi;$$

$$\delta y = -\sin \alpha \cos \theta \delta \rho + \sin \theta \delta \eta + \cos \theta \cos \alpha \delta \xi;$$

$$\delta x = \sin \alpha \sin \theta \delta \rho + \cos \theta \delta \eta - \sin \theta \cos \alpha \delta \xi;$$

thus
$$(\delta x^2) + (\delta y^2) + (\delta z^2) = (\delta \xi)^2 + (\delta \eta)^2 + (\delta \zeta)^2,$$

and the axes ξ, η, ζ are at right angles to each other.

The axes of ζ, η, ξ , however, change their direction as they move along the curve, and we have to use the equations for moving axes instead of those for axes at rest. To do this we must know the angular velocities of these axes. To begin with the axis ξ , the binormal; this axis is at right angles to the osculating plane, the reciprocal of the radius of curvature of the helix in the plane is $\sin^2 \alpha/a$, where a is the radius of the cylinder on which the spiral is wound. When the waves move through a distance

$\delta \rho = c \delta t$, the principal normal will have moved through an angle $\frac{c \delta t}{a} \sin^2 \alpha$,

hence ω_1 the angular velocity round the binormal $= c \sin^2 \alpha / a$. The reciprocal of the radius of curvature in the plane containing the binormal and ds is by the geometry of the helix $\sin \alpha \cos \alpha / a$, and hence ω_2 the angular velocity round the principal normal is $c \sin \alpha \cos \alpha / a$.

If P, Q, R are the components of a vector the rate of change of P will be

$$\frac{dP}{db} - \omega_2 R + \omega_3 Q,$$

and not dP/db .

To get the electromagnetic equations in terms of the coordinates ζ, η, ξ , we consider a rectangular parallelopiped whose sides are $d\zeta, d\eta, d\xi$. If the electric forces in these directions are R, Q, P , and the magnetic forces γ, β, α , the line integral of the magnetic forces round the boundary of the rectangle $d\xi \cdot d\eta$ is

$$\left(\frac{d\beta}{d\xi} - \frac{d\alpha}{d\eta} \right) d\xi d\eta;$$

this must equal the change in the electric flow through this area, which is

$$K \left(\frac{dR}{dt} - \omega_1 Q + \omega_2 P \right) d\xi \cdot d\eta;$$

equating these two we get

$$\left. \begin{aligned} K \left(\frac{dR}{dt} - \omega_1 Q + \omega_2 P \right) &= \frac{d\beta}{d\xi} - \frac{d\alpha}{d\eta}, \\ K \left(\frac{dQ}{dt} + \omega_1 R \right) &= \frac{d\alpha}{d\zeta} - \frac{d\gamma}{d\xi}, \\ K \left(\frac{dP}{dt} - \omega_2 R \right) &= \frac{d\gamma}{d\eta} - \frac{d\beta}{d\zeta}; \end{aligned} \right\} \dots \dots \dots (1)$$

ω_1 and ω_2 are constants and $\omega_3 = 0$.

To show that it follows from these equations that these waves can travel through space without losing energy by radiation we can apply the same method as that we use for the cylindrical waves. If an electromagnetic wave is passing through the helical circuit the expressions for the electric and magnetic quantities will contain the term $e^{i(pt - m\rho)}$. Some of these will contain this term in the form $\cos \psi$, others in the form $\sin \psi$, where $\psi = pt - m\rho$. We see from equation (1) that if R is proportional to $\cos \psi$, P, Q, α, β must be proportional to $\sin \psi$, and from the second equation that γ must be proportional to $\cos \psi$. The flow of energy is proportional to the vector product of the electric and magnetic forces; hence the flow of energy in the radial direction which is that of the principal normal is $R\alpha - P\gamma$. R and γ are proportional to $\cos \psi$ and P and α to $\sin \psi$; hence the rate of flow is proportional to $\cos \psi \sin \psi$, the average

value of which is zero. Thus helical waves would be able to pass through space without losing energy.

Since helical waves when started could pass through space without losing energy by radiation, if the electron in the cathode ray system were suddenly stopped by a collision the electronic waves might travel on and produce X-ray radiation travelling in the direction in which the electron was moving before the collision. The velocity in this direction of these X-rays would be less than the velocity of light, and their frequency when liberated would be ϕ_1 the frequency of the spin of the electron. This frequency would diminish in consequence of the Compton effect when the rays come into collision with electrons in the medium through which they were passing. This effect is an increase of wave-length by $\cdot 024 \text{ \AA}$. whatever may be the frequency of the waves.

This wave-length is the wave-length of the electronic waves before they were detached from the electron, so that a single collision would halve the frequency of the waves and destroy any resonance effects between them and an electron.

Entladungstrahlen.

Before Röntgen's discovery of X-rays E. Wiedemann had observed that when electric sparks pass through a gas, radiation, which he called *Entladungstrahlen*, is emitted, which, like ultra-violet light, has the power of producing fluorescence and thermoluminescence in certain substances when it strikes against them, is not deflected by magnetic force, and travels in straight lines.

There are many advantages, both theoretical and practical, in studying the behaviour of cathode rays when they pass through a gas at such a low pressure that they can travel a considerable distance without making a collision. I made a series of experiments on the *Entladungstrahlen* at these pressures, and found that they were produced in abundance by the stream of cathode rays *. These experiments dealt with the variation of the rays (1) with the kind of gas through which the cathode rays pass, (2) with their variation with the velocity of these rays, (3) on the part of the discharge through which the cathode rays were passing, *i. e.*, the dark space or the negative glow.

The strahlen are cut down so greatly by even the thinnest metallic foil that to filter out of them any stray cathode rays I passed them through strong electric or magnetic fields which drove out the negatively charged cathode rays but let the uncharged strahlen through. I found that some of the strahlen were so rapidly absorbed by the gas through which they were passing that they were practically all stopped by a layer of air a

* Phil. Mag. xlix. p. 761 (1925).

few millimetres thick at a pressure of $\cdot 01$ mm. of Hg. The intense absorption of a particular kind of radiation implies resonance between the frequency of the radiation and that of the absorbing body. It is shown (see p. 4) that the frequency of the electron is about $1\cdot 2 \times 10^{20}$; the wave-length of light of this frequency would be $2\cdot 5 \times 10^{-10}$ cm., or $\cdot 025$ Ångstrom units. On A. H. Compton's theory this is the amount by which the wave-length of a quantum of light is increased when it makes a collision with an electron.

The electronic waves have all the same frequency, so that the X-rays starting from them will at first have the wave-length of $\cdot 025$ Ångstrom units; these would be so intensely absorbed that except in very high vacua they would not be able to travel more than a few millimetres from the electron; when their frequency had been reduced by collisions so that they were no longer in resonance with the electron the absorption would be very much less. The loss of energy by a quantum of the easily absorbable rays need not mean the destruction of the quantum. It could be accounted for by an increase in the wave-length of the quantum by $\cdot 024$ Å.

If these views are correct X-rays furnish an extreme case of the effects produced by resonance. Some of them can be stopped by a layer of air at atmospheric pressure only a millimetre or so thick, while others can pass through several millimetres of steel.

We should expect then an X-ray *pulse* produced by the sudden stoppage or starting of an electron would not be affected by resonance to the same effect as a train of waves from a cathode ray.

Goldstein, who made very many valuable researches on the passage of electricity through gases, distinguishes between three kinds of rays—the K_1 rays (these are the ordinary cathode rays), K_2 rays which are generated from the K_1 rays, and K_3 rays produced by the K_2 rays. The K_1 -rays produce fluorescence when they strike against the glass of the tube, they start from the cathode-rays and cannot hit any part of the wall of the tube if there is a solid obstacle between it and these rays. Goldstein expresses this very happily by saying that the phosphorescence is only produced at places which have a clear view of the K_1 rays. On the view we have taken the K_2 rays are the K_1 rays which have been scattered by a collision, while the K_3 rays are the X-rays liberated in these collisions.

One of the most interesting properties of cathode rays is the production of X-rays when they strike against an obstacle. Let us consider what on the view we have taken would happen if the moving electron were suddenly stopped by a collision. Such a stoppage could not produce any effect on radiation which had been emitted by the electron before it was stopped. Hence if a sphere is described whose centre is the electron after it was stopped and whose radius is ct , where t is the interval after

the stoppage, the effect outside this sphere would be the same as if the sphere had not been stopped, *i. e.*, outside this sphere there would be a train of waves travelling in the direction in which the electron was moving before the stoppage, and having a wave-length in the direction in which the electron was moving equal to the wave-length of the electronic waves in that direction, *i. e.*, to $\cdot 025 \text{ \AA.}$; there would not be any train of waves propagated in the opposite direction. There would, however, be a spherical pulse spreading outwards with the velocity of light from the electron after it came to rest. The plane waves constitute, I think, the X-rays, but they are only a part, as they are always accompanied by a spherical *pulse* spreading outwards from the electron (see *Phil. Mag.* xlv. p. 178 *et seq.* (1898)).

On this view abrupt changes in the velocity of a cathode ray by a collision ought to produce a crop of X-rays from the electronic waves which accompanied the cathode rays before the collisions, so that X-rays may be regarded as disembodied cathode rays; they are trains of waves travelling in the direction in which the cathode ray was moving before the collision.

This makes the study of collisions between cathode rays and matter of primary importance.

If the electron in a cathode ray were stopped suddenly by a collision the helical waves it had emitted before the collision took place will travel on and have the frequency they had before the collision; as this is that of the vibration of an electron, they would be in resonance with the electrons they passed over, and so the absorption of these rays would be very great. They would, on the view taken in this paper, be X-rays of higher frequency than any but the hardest of γ -rays. Examples of the intense absorption of rays produced by the stoppage of some cathode rays are given in the paper.

The Compton effect that γ -rays colliding with an electron suffer an increase of wave-length equal to the wave-length of the electronic waves before detachment, would throw the electronic waves out of resonance with the electrons they passed over, and they would no longer be absorbed.

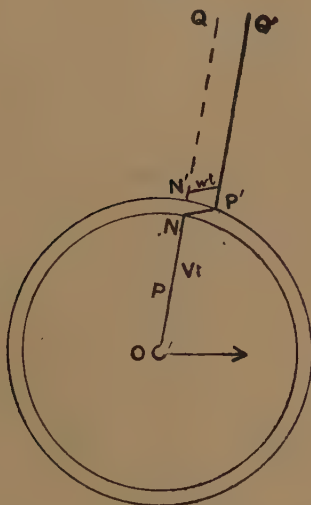
One of the most interesting properties of cathode rays is the emission of X-rays when they are stopped. On the view just put forward X-rays may be regarded as disembodied cathode rays, *i. e.*, they are cathode rays which have lost their electron.

*Pulses of Electric and Magnetic Forces produced by the
Stoppage of Charged Particles.*

The sudden stoppage of charged particles, whether the charge be positive or negative, will give rise to *pulses* of intense electric and magnetic forces. The nature of the pulse will depend on the speed of the particle before

it was stopped. If this is a small fraction of that of light the lines of electric force will be uniformly distributed round the particle, and after the particle is stopped all its lines of electric force within a certain distance of the particle must pass through it. As no electrical effects are propagated with a speed greater than that of light, after an interval t the lines of force at a distance greater than ct from the particle will be unaffected by the stoppage of the particle O . If the particle had not been stopped they would have passed through O' , the place at which the particle would have been if it had not been stopped.

Fig. 2.



Outside a thin shell at the distance ct from O the lines of electric force pass through O' , while inside this shell they must pass through O . Thus to preserve their continuity the lines of force must make a sharp bend inside the thin shell, so that they must have a tangential component while inside this shell. Each line of force while passing through the shell contributes a tangential component; the result of this is that inside this shell there is a great concentration of tangential lines of electric force. The shell spreads outwards with the velocity of light, and constitutes a spherical pulse of electric and the accompanying magnetic forces. When the velocity of the electron is small compared with c there is no plane pulse travelling in the direction in which the charge was travelling before it was stopped*.

* J. J. Thomson, 'Mathematical Theory of Electricity and Magnetism,' 5th ed. Cambridge University Press.

When, however, the charge is moving the lines of electric force before the charge is stopped are not uniformly distributed around the charge, but tend to set themselves at right angles to the direction in which it is moving, forming a thin disk at right angles to the direction of motion of the charge with radial electric and tangential magnetic forces. When the charge is stopped this distribution will move forward with the velocity of light in the direction in which it had been moving before the stoppage. All the lines of electric force have to get back to the electric charge at E inside a sphere with E as centre and radius ct where c is the velocity of light, the lines of force are straight lines passing through E . The plane sheet of electric and magnetic force has moved forward in the direction the charge was moving through a distance ct , and has left the electron E behind. The lines of electric force in passing through a thin spherical shell with centre E have to be sharply bent; this, as in the case when the velocity is small, involves an intense concentration of tangential electric force in the shell. This spherical pulse of electric force moves out radially with the velocity of light, but is now accompanied by a plane pulse moving also with the velocity of light and travelling in the direction in which the charge was moving before it was stopped.

The character of the radiation given out when electrons make collisions with matter depends on the abruptness with which the electron is stopped. Before they were stopped the radiation they emitted was the very penetrating radiation with wave-length $\lambda = 0.25 \text{ \AA.}$, and if the electron was brought to rest by the first collision it made this would be the wave-length of the X-rays produced by the impact of cathode rays. If, however, the electron were not stopped by the first collision, but had its initial velocity u reduced to $u-v$, then since the wave-length emitted by the electron is inversely proportional to its velocity, the wave-length would be $\lambda \frac{u}{u-v}$, and if it made n collisions before it was stopped the

wave-length would be increased to $\frac{\lambda u}{u-(n-1)v}$. Thus the radiation emitted before the electron was brought to rest would not be radiation of a definite wave-length corresponding to a bright line in a spectrum, but radiation corresponding to a band spectrum where the wave-length varies from λ to $\frac{\lambda u}{u-(n-1)v}$.

The study of the effects produced by collisions between cathode rays and matter would be simplified from both the theoretical and experimental point of view if the rays passed through gas at a low pressure and not through thin plates of metal.

A way by which we might hope to obtain information about the collisions between cathode rays and molecules would be to send a beam of electrons between fields of electric and magnetic forces as in the study of positive rays by the parabolic method. In this method the parabola corresponding to a given mass may have a structure which tells us much more about the particle than its mass. Thus, for example, the parabola corresponding to the atom of hydrogen sometimes has an abrupt change in darkness at a particular place, indicating that some of the atoms of hydrogen have been produced by the dissociation of a molecule.

Again, if the velocity of the moving charge were to be suddenly changed by a collision while passing through the electric and magnetic fields, the place where it struck the photographic plate would be changed abruptly, and this would produce an abrupt change in the intensity of its parabola. If it made during its passage through the electric and magnetic fields several collisions there would in consequence be a corresponding number of changes in its parabola, which would acquire a spotted appearance, and the number of spots would indicate the number of collisions.

We must distinguish between two kinds of collision between electrons and atoms or molecules. One kind of collision is one when the deflexion is due to the electron passing near the atom and being deflected by the forces between them, the deflexions might be expected to be much smaller though more frequent than those when the electron penetrates into the core of the atom. These might be expected to produce an abrupt change in the motion of the particle and change the position of the spot it makes on the photographic plate. The first type would not produce abrupt changes, but a gradual diminution in the velocities of the particles and an extension of their parabola in the region corresponding to small velocities.

We can obtain good parabolas for the electrons by using the same method as that used for the positive rays, but reversing the direction of the current through the discharge-tube, sending cathode rays through a fine tube in the anode instead of positive rays through one in the cathode.

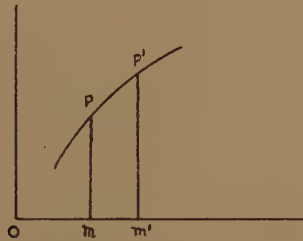
The places where the cathode rays strike against the photographic plate are all on one parabola. If the rays had all the same velocity they would strike the plate at one point, and all that would appear on the plate would be a spot at this point. If some of the rays make a collision these would again produce a spot on the parabola but not in the same place as before; thus there would be two spots on the parabola, if some had made two collisions there would be three, and so on. Thus the effect of collisions would show itself by a series of spots on the parabola due to the cathode rays. If the discharge through the gas is kept running continuously then unless all the collisions produced the same change in velocity the spots would overlap and merge into a band.

Some years ago I made experiments on electrons in connexion with another problem and obtained a number of photographs of the parabolas due to them. On examining these lately I thought I could detect traces of spots, but they were not at all well defined, and I could not feel sure that they existed. I have repeated these experiments with improved apparatus and got unmistakable evidence of the presence of a series of spots on the parabola; such a series is shown in fig. 1 on the plate*. In most of the experiments the current was produced by an induction coil; the coil was not run continuously, but a flash was sent through the discharge-tube by one "break" of the current through the coil, and photographs were taken of the parabolas due to the flash. These prove that there are groups of cathode rays with different velocities. It was important to find out whether the difference in velocity was due to collisions or whether it was due to the flash discharge not being a unit but made up of several discharges, producing different differences of potential between the electrodes of the discharge-tube and therefore different velocities in the electrons. To test this I used a discharge-tube with two aluminium electrodes of the same size and shape placed opposite to each other with their axes in the same straight line. Thus when the current through the tube was in one direction a pencil of electrons fell on the photographic plate, and when the direction of the current was reversed pencils of positively charged particles did so. The potential difference between the electrodes would be the same in the two cases, and if the spots on the electronic parabolas were due to irregularities in the potential difference they would appear on the positive ray parabolas as well as on the electronic ones. I made a great many experiments, using different gases in the discharge-tube and working over a wide range of pressure, but was never able to detect any spots on the parabolas due to the positive rays. It may be asked, why should the collisions of the positive rays with the atoms and molecules of the gas through which they pass not produce losses of velocity similar to those which occur with cathode rays?

* The figure represents the effect on the photographic plates when the cathode rays were exposed to magnetic forces only and not to a combination of both electric and magnetic forces. This was to simplify the experiments and to save space in the reproduction of the photographs. If all the cathode rays had the same velocity when they reached the magnetic field there would be but one spot in the photograph. If cathode rays with different velocities reached the plate they would produce a straight vertical line. This would be continuous if there were no abrupt changes in the velocities of the rays reaching the plates, but if there were such changes as might be produced by the rays coming into collision with the molecules of the gas through which they were passing there would be a series of spots on the vertical line, each spot corresponding to rays reaching the photographic plate with a different speed.

The answer to this question is that the experiments of Wien (*Ann. der Phys.* xxxix. p. 519 (1912)), Königsberger and Katchewski (*idem*, xxxvii. p. 101 (1912)) have shown that the positive rays do not lose an appreciable amount of energy when they pass through a gas. This was shown very simply by Königsberger and Katchewski. At two places A and B in the path of a pencil of positive rays they exposed the pencil to the action of magnetic forces, adjusted so that at a particular pressure the deflexion at B was equal and opposite to that at A, so that the particles were not deflected after passing through both magnetic fields; they found that if this balance was obtained at a particular pressure it held for all pressures. If the particles lost velocity by collision the difference between the velocities at A and B would be greater at high pressures than at low; since the deflexions depend on the velocities of the positive rays the balance would be destroyed if there were an appreciable loss of velocity in passing through the gas. Another instance of this is afforded by Deemster's

Fig. 3.



method of getting a beam of positive rays of uniform velocity. He bends by a powerful electromagnet the rays coming through one slit into a semicircle ending at another slit, and finds that the magnetic field required to do this does not depend on the pressure of the gas.

The positive rays are not positively electrified during the whole of their path; they soon attract an electron from the gas through which they are travelling and become neutral. After travelling some time in this state they lose by another collision the electron they had acquired and become positively charged again. This process goes on being repeated. Wien has made a large number of measurements of the distance they run in both the charged and uncharged state, and has shown that the free paths, though longer than those calculated from the kinetic theory of gases, are of the same order of magnitude, and that the free path of the uncharged particles is greater than that of the charged. The result of the collisions made by the positive ray particle is thus the loss or gain of an electron, and as the mass of the electron is exceedingly small compared with that of the particle the change in velocity will be very minute.

We can by measuring the distance between consecutive spots deduce the loss of energy by an electron when it makes a collision. In most of my experiments the electrons were deflected by magnetic forces alone, so that all the spots were on a straight line.

The magnetic force was adjusted so that the spot on the photographic plate made by the electrons which had not made any collisions was always brought to the same spot on the plate. If H is the magnetic force and d the distance of this spot from the spot made when the electrons were not deflected by the magnetic force

$$d = C \frac{H}{v},$$

where v is the velocity of the electron and C a constant depending only on the geometry of the apparatus. If v' is the velocity of the electron after making a collision and d' the distance of its spot from the zero

$$d' = C \cdot \frac{H}{v'},$$

so that

$$\begin{aligned} d' - d &= CH \left(\frac{1}{v'} - \frac{1}{v} \right) \\ &= d v \left(\frac{1}{v'} - \frac{1}{v} \right), \end{aligned}$$

or

$$\begin{aligned} \frac{d' - d}{d} &= v \left(\frac{1}{v'} - \frac{1}{v} \right) \\ &= \left(\frac{v' - v}{v'} \right), \end{aligned}$$

or since v' is nearly equal to v

$$\frac{d' - d}{d} = \frac{v' - v}{v}.$$

If m is the mass of the electron the change in momentum $= m(v' - v)$.

The change in energy $= \frac{1}{2}m(v^2 - v'^2)$,

so that

$$\frac{d' - d}{d} = \frac{\text{change in energy of electron}}{mv^2}.$$

Since $v' - v$ is small compared with either v or v' , this is approximately equal to $m(v' - v)/v_0$, so that we get

$$\frac{d' - d}{d} = \frac{\text{change in energy of electron}}{mv^2}.$$

The experiments indicated that $(d - d')/d$ was constant and did not depend on the potential difference between the electrodes of the discharge-tube.

This could not happen if the potential through which the electrons fell in their passage through the discharge-tube was equal to the potential difference applied to the terminals of the tube. The brightness of the spots depends greatly on the pressure of the gas, and before taking the photographs preliminary experiments were made with a fluorescent screen at the place where the photographs were to be taken and the pressure altered until the spots were brightest; this was when the cathode dark space reached nearly to the anode. In this case there is "polarization" similar to, but much greater than, that found when currents pass through liquid electrolytes. In the gas, as in the electrolyte, the current diminishes for a short time after the E.M.F. is applied to the tube, and if the direction of the current is reversed very quickly there is at first an increase in the current which is soon followed by a decrease. In some cases the decrease of the current stops after a short time, but in others, when the dark space nearly reaches the anode, the current stops altogether. The discharge is "abnormal," and so cannot persist if the current density at the cathode falls below a certain value. After a time it starts again, and then again disappears; thus the discharge through the tube is alternately bright and dark. The tube throbs, and the interval between the throbs varies enormously with slight changes in the external conditions; sometimes it is so small that the throbs cannot be detected unless they are observed after reflexion from a rapidly rotating mirror, in others the interval may be a minute or longer. It changes if the glass of the discharge-tube is touched by the finger, indicating the existence on the glass of the discharge-tube of electrification like that which produces the phenomenon of the "pseudo high vacua." The "polarization" in the tube indicates the existence of a double layer of electricity at the electrode; one of these layers is on the metal of the electrode, the other, which is inside the tube, constitutes the virtual cathode, and the effective potential difference in the tube is that between this layer and the other electrode, and not that between the leads of the battery producing the discharge, the effective potential may be much less than that applied to the tube. This was confirmed by measuring the velocity of the electrons inside the tube.

If v is the velocity of the electron, e its charge and m its mass, d the distance through which it is deflected when passing through a magnetic field, then (see 'Rays and Positive Electricity,' J. J. Thomson, p. 17)

$$d = \frac{e}{mv} H a l,$$

where l is the distance of the screen on which the deflexion is measured from the place where the magnetic force H is applied, a the distance over which H extends, the screen being so far away that l was large compared

with a . The magnetic force was that in a gap in a solenoid wound with 10 turns per cm. the current through it was 1.05 amperes.

Thus $H = 4\pi \times 10 \times 1.05/10$. Since the diameter of the solenoid was 2 cm. $Ha = 28.6$; hence

$$\frac{d}{l} = \frac{e}{mv} \times 28.6 = 1.77 \times 10^7 \times 28.6/v.$$

The deflexion of the least deflected electron, the one which had not lost velocity by collision, was 3 cm., and l was 2 cm., so that $d/l = 1/4$, and hence

$$v = 2 \times 10^9.$$

If V is the potential difference that would give this velocity to the electron

$$\frac{Ve}{m} = \frac{H}{2} \times 10^{18},$$

$$V \times 1.77 \times 70^2 = 2 \times 10^{18},$$

or, if V be measured in volts,

$$V \times 10^8 \times 1.77 \times 10^7 = 2 \times 10^{18}$$

or

$$V = 1100 \text{ volts.}$$

The potential difference applied to the electrodes of the tube was about 15,000 volts, so that the velocity of the electrons is but a small fraction of that applied to the electrodes of the tube. The distance between the electrodes of the tube was 6 cm. and the dark space was 5 cm.; so that the end of the dark spaces was not very close to the anode; had it been so the discrepancy would probably have been greater still.

The potential difference was estimated by measuring the length of the spark produced between two spheres of large radius when these were connected with the source of potential difference used to produce the discharge through the tube. This varied with the pressure of the gas, but it did not affect the loss of energy suffered by an electron at a collision. For example, in one set of experiments the equivalent spark gap varied between 6.6 and 1 mm. without producing any change in the energy lost by an electron at a collision.

The series of spots arise from different electrons having made different numbers of collisions between leaving the cathode and reaching the magnetic field where they are deflected. If the pressure of the gas is so low that the free path of the electron is large compared with the distance it has to travel between the electrodes, the chance of the electron making a collision will be small and the spot corresponding to the electrons which have not made any collision will be the brightest spot. On the other hand,

if the pressure of the gas is so high that the mean free path is small compared with the distance the electron has to travel the chance of its getting through without a collision will be very small, and so the spots for the electrons which have made many collisions will be brighter than those which have made few. There will be a certain pressure, the one when the mean free path of the electron is near the distance it has to travel, when there is little variation in the brightness of a considerable number of the spots.

In making the experiments the spots were first observed on a fluorescent screen placed where the photographic plate would be put for the photographs and the pressure of the gas altered until the spots on the screen were fairly uniform in brightness. Fig. 4 shows a case where the brightness of the spots was greater in the middle of the series than at either end. Thus it is only within narrow limits of pressure that a number of spots of nearly equal intensity can be obtained.

The series of bright spots observed on the photographs is analogous to that of bright lines in the spectrum of light, the distance between two

Fig. 4.



lines corresponding to that between two spots. From the photographs of the cathode rays we can calculate the percentage loss of energy when the electron makes a collision, and we have seen that for the same gas this is much the same for each collision. We regard the separation of the lines by equal intervals as due to the cathode rays which produce the spots having made different numbers of collisions while passing through the gas in the tube. We find from the measurement of the photographs that the percentage loss of energy at each collision is about 3.6. The sharpness of the spots shows that the collisions must be very abrupt; if they lasted for a time equal to the frequency of spin of the electron there would be "fogging" of the photograph where the electrons strike the plate, but not well-defined spots. The experiments support the view that when the spinning electrons collide with the atoms of a gas they ionize the gas and that the loss of energy at a collision is measured by its ionizing potential. The bright spots form a series whose frequencies are separated by equal intervals showing that they are due to electrons which have lost the same energy at each collision. They are to be regarded not as independent lines, but as members of a series such as the Balmer series, the number of lines being equal to the number of collisions

the cathode ray makes before reaching the place where the magnetic force is applied.

I regard all energy as being in the ether and the mass of a body as being that of the ether it can grip. The faster it moves the more it can grip and the greater its energy, while when the velocity falls it grips less, and there is more in the ether. The system behaves as if the electron had a banking account from which it drew when its velocity increased, and paid in when it diminished. The potential energy corresponds to the amount to his credit, the kinetic to the amount he has withdrawn.

We know that the mass of a body is increased when it is charged with electricity. If the body is a sphere of radius a , and moving with velocity u , then when u is small compared with the velocity of light, the energy of the sphere is $2c^2u^2/3a$. We regard the increase in energy as indicating that the charge is coupled up with the ether, and that the increase of energy is that of the ether so coupled.

The increase may be compared with the increase of the mass of a body when it moves through a liquid; it sets the liquid around it in motion, and the mass of a cylinder is increased by that of the liquid displaced, and for a sphere half this mass. If the moving body is very thin practically the whole of its mass will be due to the liquid around it.

Normal Cathode Fall of Potential.

The table on p. 25, taken from a paper by Schaufalberger (*Ann. d. Phys.* lxxiii. p. 21 (1923)), of the cathode fall of potential for the same gas with different metals for the cathode shows how great is the influence of the metal.

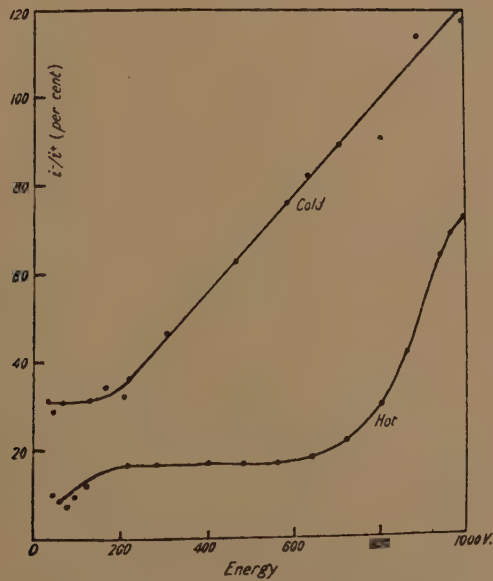
The Range of Values of the Cathode Fall with different Metals even greater than those with different Gases.

In making the nature of the gas the primary consideration it seems as if we had got hold of the wrong end of the stick. It is as if in penmanship we concentrated on the penholder and took no account of the nib. Oliphant obtained curves of the kind we have been considering only after long and continuous heat treatment—indeed, with some metals, aluminium, nickel, and tungsten, he never obtained them at all. The curve which is obtained if the heat treatment is not so drastic as that used by Oliphant in his final experiments is represented by the straight line marked “cold” in fig. 5. In this the relation between the number of electrons emitted by the impact of a positive ion and the energy of the ion is represented by a straight line, and is always greater than for the more perfectly gassed electrodes. This type is more capricious than the other, it cannot be reproduced with certainty. We cannot regard the cathode fall as fixed even by the metal of the cathode; it would, I think, be nearer the truth to say that each cathode has a cathode fall of its own

which cannot be got from tables, but must be determined by an *ad hoc* experiment.

Cathode metal.	Dark space (in cm.).	Cathode fall.
Mg	·614	247
Al	·669	302
Be	339
Ni	·893	353
Fe	·880	363
Zn	·807	372
Cd	·869	375
Cu	·893	375
Ir	379
Co	381
Pb	·839	392
Sn	393
Sb	396
Au	418
Pd	421
Pt	1·026	425
Aq	428

Fig. 5.



When the conditions in the dark space and negative glow are steady then for any positive atom which strikes against the cathode another

must come from these regions. Electrons coming from the cathode and passing through the dark space might ionize an atom, and thus give it a positive charge and lose some of their energy in the process. Some might make more than one collision, so that some electrons would lose more energy, and when acted upon by magnetic force would be differently deflected. If before the magnetic force was applied they struck a fluorescent screen at the same spot the magnetic force would draw the spot out into a band. It is one of the most interesting things about the electric discharge through gases that this does not occur, the spot is as well defined as before. This indicates that ionization in the dark space is not sufficient to supply the loss of the positive atoms which have struck against the cathode—indeed, there is no direct evidence that it does anything appreciable towards it. In ‘Discharge of Electricity through Gases’ an estimate is given of the loss of energy in the dark space by ionization, making various assumptions as to the alteration of ionization with the velocity of the electron, but at most this amounted to a very small percentage of the initial energy of the electron. We must therefore look to the negative glow to supply the positive ion required to sustain the discharge.

After the electron has left the dark space the electric force is so small that when it loses energy by ionizing an atom or molecule there is no force acting on it to increase its energy, and so this will soon be lost. We shall have now since an atom has lost an electron, (1) a positively electrified atom, (2) the electron taken from it when it was ionized, (3) the original electron with diminished energy.

The experiments made on the mobility of ions prove that the mobility of the negative ion shows a remarkable increase as compared with that of the positive when the pressure is low, and in ‘Recent Researches’ reasons are given which seem convincing for supposing that this is due to the negative ion beginning as an electron but after a time getting attached to a neutral atom, and producing one with a negative charge—so that the negative glow could supply negatively as well as positively electrified ions.

As this potential was increased the rate of increase of the emission diminished, and at 1000 volts showed distinct signs of approaching a limiting value. If we take the voltage at which it reached this value as 1200 volts, and the loss of energy at a collision as that due to the emission of an electron from the helium atom, and measured by 25 volts, by the formula on p. 9,

$$\frac{d_1 - d_2}{d} = \frac{25}{1200} = 2 \text{ per cent.},$$

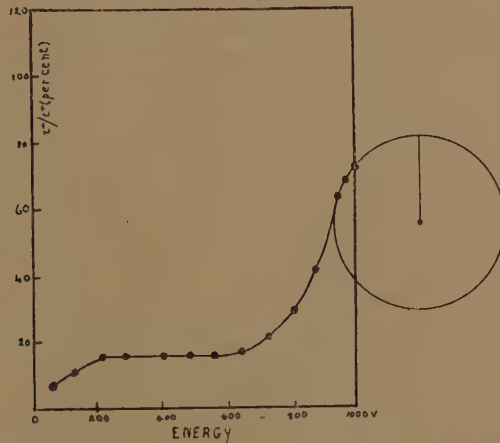
which is about the value found in my experiments on helium.

The limitation of the velocity of atoms in a discharge-tube is due

to the distribution of electric force which exists in a gas at low pressure through which a discharge is passing.

The estimate of 1200 volts was arrived at by drawing a circle through the last three points on the curve; if the centre of the circle is at O then at a point vertically above O the curve would be horizontal and there would be no change in the number of electrons emitted as the potential difference increased. We see from the curve that the emission is not appreciable until the energy of the positive ion is as great as 100 volts; after this it increased rapidly with the voltage until this reached 600. There was little increase between 600 and 800 volts, when it began again to increase, and reached a maximum value at 1200 volts. Since an increase in the voltage of 200 volts produced no increase in the emission of electrons the earlier observers may have come to the conclusion that they had

Fig. 6.



reached the limit of the velocity which could be attained by a positive ion. The emission of electrons by the impact of positive ions is a very complicated phenomenon. Layers of gas over the surface produce great effects, impurities in the gas from which the positive atoms are produced also exert a great influence over the magnitude of the effect. The most thoroughgoing experiments which have been made on this subject are those of Professor Oliphant (Proc. Roy. Soc. cxxvii. p. 373 (1930)), using very carefully purified helium, and taking great precaution to free the surfaces of the electrodes from layers of gas. He found very few electrons were emitted until the energy of the helium atoms corresponded to a potential fall of about 100 volts; this is about the maximum value which, according to Frank's experiments on the Döppler effect, could be acquired by a charged particle. At this voltage the number of electrons rose abruptly.

So that in the negative glow there are, in addition to the uncharged particles, positively and negatively electrified ions, and in gases very carefully freed from electronegative elements free electrons * abruptly, then for a considerable range of velocities the increase in emission was proportional to the increase in the potential difference through which the positively charged atoms fell.

Abnormal Discharge.

In the "abnormal discharge" the cathode fall of potential is no longer constant, the essential condition is that the mean current density at the cathode should not fall below a definite value. This depends upon what occurs at the junction of the metal and the gas, and there are those which experience shows are most capricious and to be affected by minute quantities of impurities in the gas or minute changes in the condition of the surface of the electrodes. They cannot be got from tables of physical constants, but must be determined for each experiment. This was done in the experiments of Professor Oliphant (Proc. Roy. Soc. cxxvii. p. 379), and it is for that reason that I think that the close agreement between these experiments and those indicated by the theory I have put forward is the strongest argument I know in favour of that theory.

In order that the discharge should be self sustaining each positive ion which strikes against the cathode must give rise to an electron travelling in the opposite direction, and since its charge is negative producing the same current as the positive. Thus the problem of the emission of electrons by the impact of positive ions investigated by Professor Oliphant is of primary importance in the theory of the discharge. This emission was not appreciable until the energy was that of 100 volts, when it very rapidly increased; it remained fairly constant up to potential differences between the electrodes of 600 volts, when it began to increase; at 800 volts there were distinct signs that the rate of increase with the potential difference diminished and at 1000 volts the curve looked as if the limit had been nearly reached. Since between 200 and 700 volts there is hardly any increase in the number of ions emitted from the metal plate, it is possible that the earlier experimenters may have thought that with 100 volts the maximum velocity had been reached, and that no greater velocity could be acquired by a charged atom.

His experiments indicate that with low pressures and high potential differences between the electrodes the potential between the electrodes is represented by a curve like that in fig. 5, where the ordinates represent the potential at a distance from the cathode represented by the abscissæ.

* Evidence in favour of the existence of negative *ions* in the negative glow is given on p. 31.

I showed (see Phil. Mag. xviii. p. 461 (1909)) that in the electric discharge through gases at low pressures the electric force near the anode was in the opposite direction to the direction of the current. An extreme case of currents with small difference of potentials between the terminals was one where a beam of cathode rays was sent through a small hole in a metallic box where a discharge with negative dark space negative glow though these were there.

Another illustration of the large effects produced in the discharge of electricity through gases between metallic electrodes by adventitious causes is afforded by some experiments I made (Phil. Mag. xi. p. 698 (1931)), when the dark space reached very nearly to the anode, then, though the electrodes were two circular plates of the same size and made of the same metal, the discharge might pass through the tube in one direction but not in the opposite. The discharge showed polarization effects similar to those found in liquid electrolytes. When first the potential was applied the current began to decrease. Sometimes this decrease would stop after a time, in others it went on until the current stopped altogether. After an interval it started again, went up to about its old value, and then began to decrease again and the process was repeated. Thus the tube kept on lighting up and then going out. The interval between the "throbs" varied enormously—in some cases they followed each other so quickly that a rotating mirror was required to detect them, in others there might be a minute or more between the throbs. When a pad of cotton-wool soaked in liquid air was pressed against the outside of the discharge-tube the brightness of the discharge after an interval began to diminish and sometimes fell to a small fraction of its original value. When the pad was removed the brightness of the discharge began to increase slowly at first, but afterwards very rapidly, rising far beyond that before the application of the pad. After a time it fell to the value it had before the pad was applied. The pad froze out of the tube some volatile gas, water vapour probably which greatly increased the discharge, and experiments on the discharge would be measurements of the quantity of vapour, and not on the electrical properties of the gas in the tube.

When the pressure of the gas is so low that the dark space is nearly equal to the distance between the electrodes in the discharge-tube, and therefore the difference in potential between them very great, the parabolas for the electron sometimes show a prolongation in the *opposite* direction to the displacement of the spots, indicating that some of the electrons have their velocity increased when they pass through the gas; the prolongations themselves show spots, indicating that some of the electrons producing it had lost momentum by collisions. In general when the prolongation occurred the pressure in the discharge-tube diminished, indicating that some chemical combination was going on. This generally

stopped after a short time, and when it stopped the prolongation disappeared. We have mentioned before the great effect produced by small quantities of impurities in some cases. A whiff of benzene vapour introduced into the tube when the discharge was passing at once produced this prolongation to a very marked extent when the electrodes were made of aluminium, but did not do so if the electrodes were made of graphite. This suggests that the effect was due to some chemical action at the electrodes.

Although there are a series of spots separated by equal intervals when the pressure of the gas through which the rays are passing lies between certain limits, there are frequently spots present which do not obey this rule, and these make the spacing between adjacent spots appear irregular; they seem more crowded together in some places and more widely separated in others. There are so many things that might influence the position of the spots that the wonder is not that there should sometimes be irregularities but that there should ever be regularity.

Among the causes which might be expected to produce irregularities in the distribution of the spots is the existence of cathode rays proceeding from different sources, and it is only the plates where the spots are regularly spaced that I have used for measurements.

The sharpness of the spots shows that *abrupt* changes in the momentum of the cathode rays occurs and that these changes are of approximately equal magnitude. This is what we might expect from what we know about the properties of cathode rays. These suffer deflexions of two kinds—one when they strike against an electron and eject it from an atom or molecule. This will cause the electron in the ray to lose *abruptly* an amount of energy measured by the ionization potential of the atom. Other collisions which do not result in ionization will produce a deflexion, though not an abrupt one, and will not produce a sharp spot on the photograph, but a diffuse fogging of the plate. The loss of energy is measured by the ionization potential of the gas, and a good many gases have potentials which are not very different. This is shown by the following table:—

Ionization Potentials in Volts.

Arg	15.4	Ne	21.47
H	13.54	O	13.6
He	24.48	C ₆ H ₆	9.6
Hg	10.31	H ₂ O	13.0
N	14.5		

The greatest differences are between normal and inert gases, and in my experiments I found that the results with helium 98 per cent. pure were irregular. Some of them gave 2 per cent. as the percentage loss of energy at a collision, others gave the same values as the ordinary gases.

When a thin beam of cathode rays is sent through a long tube containing gas at a low pressure the beam, though sharply defined where it starts, as it travels further begins to get diffuse and be accompanied by a glow which gets more and more developed as the beam travels on until the tube is filled with glow. If the beam is deflected by a magnet the glow will accompany it, even when it is twisted into spirals with many turns.

I think this is due to electrons in the beam becoming attached to the atoms of the gas through which they are passing and producing a negatively charged atom, and that it is this atom which gives out the glow.

The glow is most strongly developed in electro-negative gases like oxygen, where the passage between the dark space and the negative glow is very sharply defined, while it is very difficult to locate in carefully purified helium. The experiments of Frank and Herz showed that in carefully purified argon, electrons could make many thousands of collisions before becoming attached to an atom. The existence of negatively charged atoms of many different kinds is a well known phenomenon in positive ray analysis. The glow is absent from the dark space, but in it the electrons are moving with a very high velocity, and though an atom might be able to keep a negative charge it could not capture a fast one. As soon as the electrons get out of the dark space and lose their velocity by collisions they are captured by the atoms of the gas and give out the negative glow.

There is very interesting evidence from experiments that there are negatively charged atoms in the negative glow, interesting because ten years before I made my experiments on the electron Sir Arthur Schuster had applied the magnetic deflexion of moving charged bodies to determine the mass of the negatively charged particles in the negative glow, and found that it was equal to the mass of an *atom* of the gas. He assumed that the charged particles had made so many collisions that they had lost the velocity they had acquired under the electric field near the cathode, and that their energy corresponded to that due to thermal agitation. This is approximately the case for the charged *atoms* in the negative glow, and we may regard his experiments as a direct proof of their existence.

I am greatly indebted to my son, Professor G. P. Thomson, F.R.S., for his criticisms in the discussions I have had with him on electronic waves, but I wish to make it quite clear that this does not imply that he is in any way committed to the views about electronic waves expressed in this paper. Mr. Morley, my assistant, has given most valuable help; he has taken all the photographs and made some valuable improvements in the apparatus.

SUMMARY.

The experiments made on electronic waves have shown that if u is the velocity of the electron in the direction in which it is travelling and λ the wave-length of the electronic wave $u\lambda$ is constant and equal to $2\pi c^2/p$ where p is the frequency of the electronic wave; it follows from this equation that all electronic waves have the same frequency and naturally suggests that this is the frequency of the electronic spin, and that this is 1.2×10^{20} . Though u is in general very much smaller than the velocity of light, yet the wave and the electron keep in touch with each other; to do this the wave must pursue a very circuitous path. The case when this path is a regular helix is worked out, the velocity of the wave along the path of the spiral is $c \cos \alpha$, if this equals u the two will keep together. It is shown that the electron does not lose energy by radiation as it travels through space. A large number of experiments have been made on the loss of energy by collisions between cathode rays and the gas through which they are passing, using the parabolic method of crossed electric and magnetic fields. Since all the electrons have the same mass there is only one parabola, and if they all had the same speed there would only be one point; but if the electron loses speed by an abrupt collision there will be another spot, if it made two there would be three, and so on. The collision and the changes produced in velocity are registered on the photographic plate. It was found that the changes of velocity were due to the ionization of the gas by collisions and that the energy loss was measured by the ionizing potential of the gas. This in good accord with the experiment. A most important point is that it is not an isolated spectrum; it is a series spectrum, the lines are not independent, each is a part of a whole, each line is accompanied by a number of others whose frequencies form a series with like Balmer's Series the same difference between neighbours.

In order to keep in equilibrium with its surroundings an electron must have the energy possessed by one whose frequency is that of the spinning electron, there is a continuous supply of energy of this frequency and of no other, when things are in a steady state each electron will have bound to it this amount of energy in the ether.

II. *The Concept and Determination of Mass in Newtonian Mechanics.*

By V. V. NARLIKAR, Benares Hindu University, India *.

[Received August 29, 1938.]

ABSTRACT.

Following Mach's exposition, Dr. Pendse has recently pointed out two serious lacunæ in the definition and determination of mass. In this paper we show that a more penetrating analysis of Newtonian gravitation on the lines suggested by Mach himself and followed by Einstein, Eddington and Milne, in their respective theories, removes the lacunæ. The definition and determination of mass are thus shown to be unambiguous in the classical theory.

1. IN a recent number of 'The Philosophical Magazine,' Dr. C. G. Pendse⁽¹⁾ has written a very interesting paper entitled "A Note on the Definition and Determination of Mass in Newtonian Mechanics." Following Mach's⁽²⁾ exposition of the concept of mass, substantially based on Newton's Second and Third Laws of Motion, Pendse obtains the striking result that if Mach's procedure is applied to a situation involving more than two gravitating particles the masses of the particles need not be determinate; and, in fact, he has proved that if there are more than seven particles in the field any number of observations of the accelerations will be insufficient to determine the mass of each. Mach, himself, was aware of a weakness in his exposition of the concept of mass: he could not account for the empirical fact that the mass of a particle A is the same, whatever be the other particle, the interaction with which decides the mass of A. He therefore included the following proposition⁽³⁾ in his formulation of the foundations of mechanics:

"Experimental Proposition: The mass-ratios of bodies are independent of the character of the physical states (of the bodies) that condition the mutual accelerations produced, be those states electrical, magnetic or what not; and they remain, moreover, the same whether they are mediately or immediately arrived at." Pendse has also pointed out that a postulate, equivalent in its content to the above proposition, must be incorporated in the Newtonian system. Thus there is a lacuna in the

* Communicated by the Author.

definition of mass, and Mach made up for it by his experimental proposition; there is a lacuna also in Mach's method of determination of mass as Pendse has demonstrated.

Mach certainly elucidated the concept of mass by defining the mass-ratio of two gravitating particles as the negative reciprocal of the accelerations set up in them. It is doubtful, however, whether he opined that this method could be extended to determine the mass-ratios of n gravitating particles. We have not come across any explicit statement of his to that effect. Karl Pearson⁽⁴⁾ has, however, stated quite explicitly that the masses of particles are determined, in practice, in this manner; and Pendse's calculations prove beyond doubt that this statement of Pearson's is wrong.

It does not seem plausible that a complete knowledge of the masses in a system of n gravitating particles can, in general, be acquired from the Second and Third Laws of Motion—that is, without using the law of gravitation and the law of inertia. That these lacunæ should be discovered is not therefore surprising. We will now proceed to show how they can be removed after a critical examination of the circumstances of Newtonian gravitation.

2. Consider a gravitational situation in which n bodies are involved. The celestial bodies present many such situations. If for all relevant observational purposes the structure of the bodies can be neglected we can treat them as idealized particles. This step is consistent with the deduction from D'Alembert's principle that the motion of rotation of a body can be dynamically (as well as kinematically) accounted for independently of its motion of translation. Now one of the main problems of Newtonian Mechanics is to understand, with a view to predict them, the motions of the n idealized particles. With a formal solution of this problem one arrives not only at the definition of mass, but also at a method of determining the mass of a particle.

The motions of the n particles can be best understood in classical terms by postulating a double influence on each: one due to all the remote bodies of the universe and another influence due to the remaining $(n-1)$ particles of the system. The first influence is that of the substratum of the universe, and it is partly contained in the First Law of Motion. The second influence is well accounted for in the Newtonian law of gravitation. Even in Milne's⁽⁵⁾ theory of gravitation, where the inverse-square law is deduced, gravitation is shown as the local influence superposed on that of the substratum. In general relativity as well there are two influences at work in any gravitational situation: the local influence is revealed in the curvatures of space-time and the influence of the substratum is found in the boundary condition at infinity, viz., that the space-time is flat there.

3. Mach was the first to recognize the nature of the influence that the substratum exercises. If there were only one body in the universe its mass-property would not have been thought of, and if there were only two bodies in the universe they would not have attracted each other according to the inverse-square law : these were Mach's views. Sir Arthur Eddington⁽⁶⁾ has adopted similar views in interpreting a term of the equation of the hydrogen atom. He has argued that the equation of the hydrogen atom (one-proton, one-electron system) must contain a term which gives the influence of the substratum of the Universe. In Newtonian mechanics, we suggest that the influence of the substratum is given by the equations

$$(A) \frac{du}{dt} = 0, \quad (B) \frac{dm}{dt} = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and not by just

$$\frac{d}{dt}(mu) = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

as it is usually stated. The equation (2) or (A) is interpreted as the First Law of Motion, u being the velocity of the influenced particle at any instant. The equation (2) is really contained in the Second Law of Motion. The novelty of our suggestion lies in the equation (B), which means that with each particle is associated an unchanging number m . The number, m , is called the mass of the particle. In the general theory of relativity⁽⁷⁾ the analogue of (2) is given by

$$\frac{d}{ds} \left\{ m \frac{dx^\alpha}{ds} \right\} + m \left[\begin{smallmatrix} \mu \\ \alpha\beta \end{smallmatrix} \right] \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0, \quad . \quad . \quad . \quad . \quad (3)$$

which splits up of itself into

$$(A') \frac{d^2 x^\mu}{ds^2} + \left[\begin{smallmatrix} \mu \\ \alpha\beta \end{smallmatrix} \right] \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0, \quad (B') \frac{dm}{ds} = 0, \quad . \quad . \quad . \quad . \quad (4)$$

(A'), (B') are the analogues of (A) and (B) respectively. But while (A'), (B') follow from (3), (A), (B) do not follow from (2). Hence our explicit suggestion as contained in (1).

In the light of the equation (B) we now understand the reason why Mach's experimental proposition, referred to above, is true. There is no lacuna now in the definition of mass. Mach's procedure is ineffective for the actual determination of the invariant mass-numbers. But the inverse-square law can be used for this purpose. A question arises at this stage which cannot be left unanswered. If the mass-numbers are fixed by the substratum, will they not change if the substratum itself undergoes a motion as in an expanding universe? Einstein⁽⁸⁾ conjectured from Mach's writings that the mass-number of a particle will diminish if all the bodies constituting the substratum recede from it. He actually

verified that this is true in his theory by considering the influence of the motion of ponderable bodies on the field of an isolated particle.

For the determination of the masses of the particles we consider the local influence which is summed up in the inverse-square law of gravitation. Let the n unknown masses be m_1, m_2, \dots, m_n and let m_r be at x, y_r, z_r at time t . If the acceleration of m_1 along ox is X , at time t ,

$$X = \frac{m_2(x_2 - x_1)}{r_{12}^3} + \frac{m_3(x_3 - x_1)}{r_{13}^3} + \dots + \frac{m_n(x_n - x_1)}{r_{1n}^3}. \quad (5)$$

Here r_{12} is the distance between m_1 and m_2 , r_{13} , r_{14} , etc., having similar meanings. All the distances involved in (5) can be observed and so also the acceleration, X . Thus we have an equation of the form

$$A_2 m_2 + A_3 m_3 + \dots + A_n m_n = X, \quad (6)$$

where A 's and X are known. We can get $(n-1)$ independent equations like (6) from observations at $(n-1)$ instants. Thus $m_2, m_3 \dots m_n$ become known in terms of the observables. We need consider similarly an acceleration-component of any other mass to get the value of m_1 . All the masses become known in this manner.

References.

- (1) C. G. Pendse, "A Note on the Definition and Determination of Mass, etc." *Phil. Mag.* (7) xxiv. pp. 1012-1022 (1937).
- (2) E. Mach, 'The Science of Mechanics,' pp. 216-225 (1893).
- (3) E. Mach, *loc. cit.* p. 243.
- (4) Karl Pearson, 'The Grammar of Science,' p. 303 (1899).
- (5) E. A. Milne, "On the Foundation of Dynamics." *Proc. Roy. Soc. A*, cliv. pp. 22-52 (1935). The reader may also be referred to the more recent papers by Milne in the same journal.
- (6) A. S. Eddington, 'The Expanding Universe,' p. 111 (1933). Also 'Relativity Theory of Protons and Electrons,' p. 153 (1936).
- (7) A. S. Eddington, 'The Mathematical Theory of Relativity,' p. 127 (1924).
- (8) A. Einstein, 'The Meaning of Relativity,' pp. 110-113 (1922).

III. *The Flexure of an Infinite Elastic Strip on an Elastic Foundation.*

By GEOFFREY BOSSON, M.Sc.*

[Received July 4, 1938.]

Introduction.

MR. R. THORNTON COE in a recent paper⁽¹⁾ gave in some detail an interesting elementary treatment of the bending of beams on elastic foundations, and the present writer felt that it would be interesting to investigate the extent to which the assumptions of the elementary theory are upheld by a more exact analysis. The method employed in the present paper is that of "Generalized Plane Stress"⁽²⁾ and, to simplify the analysis, the beam is supposed to be of infinite length. This enables Fourier Integrals to be used instead of (as would be necessary in the case of a beam of finite length) Fourier Series. Much of the work would, however, be the same in the case of a beam of finite length, and the appropriate changes from integrals to series could readily be made.

We consider a strip of rectangular cross-section bounded by

$$x = \pm \infty, \quad y = 0, \quad y = b, \quad z = \pm \frac{c'}{2}.$$

The mean tensions $\bar{x}\bar{x}$, $\bar{y}\bar{y}$ and the mean shear $\bar{x}\bar{y}$ are then obtainable from a "Stress function," $\chi(x, y)$, which satisfies the biharmonic equation

$$\nabla^4 \chi = 0 \quad \left(\text{where } \nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$$

by means of the equations

$$\bar{x}\bar{x} = \frac{\partial^2 \chi}{\partial y^2},$$

$$\bar{y}\bar{y} = \frac{\partial^2 \chi}{\partial x^2},$$

$$\bar{x}\bar{y} = - \frac{\partial^2 \chi}{\partial x \partial y}.$$

The mean displacements \bar{u} , \bar{v} are then given by

$$E\bar{u} = \frac{\partial \psi}{\partial y} - (1 + \eta) \frac{\partial \chi}{\partial x},$$

$$E\bar{v} = \frac{\partial \psi}{\partial x} - (1 + \eta) \frac{\partial \chi}{\partial y},$$

* Communicated by the Author.

By expressing the condition that χ' is bi-harmonic we obtain, after some reduction, the symbolic form⁽⁴⁾

$$\chi' = (\cos yD + \frac{1}{2}yD \sin yD)X_0 + \left(\frac{3}{2D} \sin yD - \frac{1}{2}y \cos yD\right)X_1 \\ + \frac{y}{2D} \sin yD \cdot X_2 + \frac{1}{2}\left(\frac{\sin yD}{D^3} - \frac{y}{D^2} \cos yD\right)X_3. \quad (2)$$

where

$$D \equiv \frac{d}{dx}.$$

The corresponding displacement function, ψ' , is given by

$$\psi' = \sin yD \cdot X_0 - \frac{1}{D} \cos yD \cdot X_1 + \frac{1}{D^2} \sin yD \cdot X_2 - \frac{1}{D^3} \cos yD \cdot X_3. \quad (3)$$

Further, we find that

$$(\widetilde{y}y)_{y=0} = D^2 X_0 \quad (4)$$

and

$$(\widetilde{x}y)_{y=0} = -DX_1. \quad (5)$$

Now on our hypothesis $(\widetilde{x}y)_{y=0} = 0$, and we may therefore take $X_1 = 0$, since, from equation (5), X_1 can at most be a constant and would not as a consequence of this contribute anything to the stresses.

Returning to (2), we have

$$-\widetilde{x}y = \frac{1}{2}(yD^3 \cos yD - D^2 \sin yD)X_0 \\ + \frac{1}{2}(\sin yD + yD \cos yD)X_2 + \frac{1}{2}y \sin yD \cdot X_3. \quad (6)$$

Accordingly $y = -f$ is free from shear if

$$(fD^3 \cos fD - D^2 \sin fD)X_0 + (\sin fD + fD \cos fD)X_2 - f \sin fD \cdot X_3 = 0. \quad (7)$$

Now

$$E'\bar{v} = \frac{\partial \psi'}{\partial x} - (1 + \eta') \frac{\partial \chi'}{\partial y} \\ = D \sin yD \cdot X_0 + \frac{1}{D} \sin yD \cdot X_2 - \frac{1}{D^2} \cos yD \cdot X_3 \\ - \frac{(1 + \eta')}{D} [\frac{1}{2}(yD^3 \cos yD - D^2 \sin yD)X_0 \\ + \frac{1}{2}(\sin yD + yD \cos yD)X_2 + \frac{1}{2}y \sin yD \cdot X_3], \quad (8)$$

and therefore

$$E'(\bar{v})_{y=0} = -\frac{1}{D^2} \cdot X_3. \quad (9)$$

Again, from (8), $\bar{v} = 0$ on $y = -f$ if

$$D \sin fD \cdot X_0 + \frac{1}{D} \sin fD \cdot X_2 + \frac{1}{D^2} \cos fD \cdot X_3 = 0, \quad (10)$$

since the part of (8) included in square brackets vanishes when $y = -j$ in virtue of (7).

As regards the effect of the foundation on the strip, we need only concern ourselves with the functions X_0 and X_3 , which, in virtue of (4) and (9), are the only ones which affect the values of $(\bar{y}y)_{y=0}$ and $(v)_{y=0}$. The desired condition is therefore obtained by eliminating X_2 between (7) and (10). This gives ultimately

$$-2D \cdot \sin^2 fD \cdot X_0 = \left(\frac{1}{D^2} \sin fD \cdot \cos fD + \frac{f}{D} \right) X_3, \quad \dots \quad (11)$$

that is

$$2 \sin^2 fD \cdot (\bar{y}y)_{y=0} = \left(\frac{1}{D} \sin fD \cdot \cos fD + f \right) D^2 E' (v)_{y=0}. \quad \dots \quad (12)$$

If we assume that the loading on the edge $y=b$ of the strip is such as to keep the strip in contact with the foundation, then \bar{v} and $\bar{y}y$ will be continuous in crossing the interface. It follows, therefore, that the stress function for the strip must be such that it gives a value of $(\bar{y}y)_{y=0}$ and a value of $(v)_{y=0}$ which satisfy equation (12).

The Stress Function for the Strip.

It is now necessary to obtain a stress function which will satisfy equation (12), give zero shear on $y=0$ and $y=b$, and give an assigned distribution of load on $y=b$. We divide the problem into two cases according as the distribution of load is even or odd. The general case is then easily obtained by combination.

In the former case, we take a stress function χ given by

$$\chi = F(y) \cos mx, \quad \dots \quad (13)$$

where

$$F(y) = A \cosh my + B y \sinh my + C \sinh my - Cmy \cosh my. \quad \dots \quad (14)$$

This, as it stands, gives zero shear on $y=0$. The condition for zero shear on $y=b$ is that

$$Ams + B(s + mbc) - Cm^2bs = 0, \quad \dots \quad (15)$$

where

$$s = \sinh mb \quad \text{and} \quad c = \cosh mb.$$

The displacement function, ψ , corresponding to χ , is readily found to be given by

$$\psi = \frac{2}{m} (B \sinh my - Cm \cosh my) \sin mx. \quad \dots \quad (16)$$

Now

$$E\bar{v} = \frac{\partial \psi}{\partial x} - (1 + \eta) \frac{\partial \chi}{\partial y},$$

and therefore

$$E(\bar{v})_{y=0} = -2Cm \cos mx, \quad . \quad . \quad . \quad . \quad . \quad (17)$$

while, corresponding to χ , we have

$$(\bar{y}y)_{y=0} = -m^2A \cos mx, \quad . \quad . \quad . \quad . \quad . \quad (18)$$

and if we substitute these values into (12) we obtain

$$2 \sinh^2 mf \cdot A = \frac{E'}{E} (\sinh 2mf + 2mf)C. \quad . \quad . \quad . \quad . \quad (19)$$

Let us now adopt the following abbreviations :

$$\frac{E'}{E} = \lambda, \quad 2 \sinh^2 mf = \sigma \quad \text{and} \quad \sinh 2mf + 2mf = \sigma',$$

then (19) becomes

$$\sigma A - \lambda \sigma' C = 0, \quad . \quad . \quad . \quad . \quad . \quad (20)$$

and we have from (15) and (20)

$$\frac{A}{\lambda \sigma' (s + mbc)} = \frac{B}{(\sigma mb - \lambda \sigma') ms} = \frac{C}{\sigma (s + mbc)}. \quad . \quad . \quad . \quad . \quad (21)$$

We may therefore take

$$\begin{aligned} \chi = & [\lambda \sigma' (s + mbc) \cosh my + (\sigma mb - \lambda \sigma') msy \sinh my + \sigma (s + mbc) \sinh my \\ & - \sigma (s + mbc) my \cosh my] \cos mx, \quad . \quad . \quad . \quad . \quad (22) \end{aligned}$$

and this may be generalized, subject to suitable restrictions of $f(m)$, by taking a new stress function χ_0 given by

$$\chi_0 = \int_0^\infty \chi \cdot f(m) \cdot dm. \quad . \quad . \quad . \quad . \quad . \quad (23)$$

Employing now the new stress function, χ_0 , we have formally

$$(\bar{y}y)_{y=b} = - \int_0^\infty [\lambda \sigma' (sc + mb) + \sigma (s^2 - m^2 b^2)] m^2 f(m) \cos mx \cdot dm,$$

and so, if the load per unit area applied on $y=b$ is an even function, $W(x)$, we shall have $(\bar{y}y)_{y=b} = -W(x)$, and therefore

$$W(x) = \int_0^\infty [\lambda \sigma' (sc + mb) + \sigma (s^2 - m^2 b^2)] m^2 f(m) \cos mx \cdot dm.$$

It now follows from Fourier's reciprocal integral theorem for even functions that

$$f(m) = \frac{2}{\pi} \times \frac{1}{m^2 [\lambda \sigma' (sc + mb) + \sigma (s^2 - m^2 b^2)]} \int_0^\infty W(\xi) \cos m\xi \cdot d\xi. \quad . \quad (24)$$

This value of $f(m)$ is now substituted into (23) to give the general stress function for the strip in its final form :

$$\chi_0 = \frac{2}{\pi} \int_0^\infty \lambda \sigma' (s + mbc) \cosh my + (\sigma mb - \lambda \sigma') msy \sinh my \\ + \frac{\sigma (s + mbc) \sinh my - \sigma (s + mbc) my \cosh my}{m^2 [\lambda \sigma' (sc + mb) + \sigma (s^2 - m^2 b^2)]} \\ \times \cos mx . dm \int_0^\infty W(\xi) \cos m\xi d\xi . \quad (25)$$

Special Case when f is Large.

If f is large, equation (19) becomes ultimately

$$A = \lambda C \quad (25 a)$$

and (25) then reduces to

$$\chi_0 = \frac{2}{\pi} \int_0^\infty \lambda (s + mbc) \cosh my + (mb - \lambda) msy \sinh my \\ + \frac{(s + mbc) \sinh my - (s + mbc) my \cosh my}{m^2 [\lambda (sc + mb) + s^2 - m^2 b^2]} \\ \times \cos mx . dm \int_0^\infty W(\xi) \cos m\xi d\xi . \quad (26)$$

It is easy to show that, if $W(x)$ is finite and integrable over the range $0, \infty$, then the integral (25) converges uniformly at the upper limit if $y < b - \epsilon$.

In the region of $m=0$ we have

$$\text{Integrand} = \frac{8\lambda fb [1 + \frac{1}{3}(f^2 + b^2)m^2 + \dots] \cos mx}{8\lambda fb m^2 + \dots} \int_0^\infty W(\xi) . \cos m\xi . d\xi . \quad (27)$$

Now if $\int_0^\infty W(\xi) . d\xi$ exists and is finite, then

$$\left| \int_0^\infty W(\xi) \cos m\xi . d\xi \right| \leq \left| \int_0^\infty W(\xi) . d\xi \right| < K,$$

where K is some positive constant independent of m . It follows at once from (27) that, provided $\lambda \neq 0$, the only divergent part of the integrand is

$$\frac{1}{m^2} \int_0^\infty W(\xi) \cos m\xi . d\xi,$$

and this, being independent of x and y , will not contribute to the stresses and may, in fact, be removed by adding a term such as

$$-\frac{1}{m^2} . e^{-ms} \int_0^\infty W(\xi) \cos m\xi . d\xi$$

without affecting either the stresses or the convergence at infinity. We shall consequently treat (25) as though it were a convergent integral.

Differentiations with respect to x or y will not affect the convergence at the upper limit and, since they will introduce additional powers of m into the numerator, they will merely render the convergence at the lower limit more rapid.

The Displacement Function.

The displacement function ψ_0 , corresponding to (25), is readily found to be

$$\psi_0 = \frac{4}{\pi} \int_0^\infty \frac{(\sigma mb - \lambda \sigma') s \sinh my - \sigma(s + mbc) \cosh my}{m^2[\lambda \sigma'(sc + mb) + \sigma(s^2 - m^2 b^2)]} \times \sin mx \cdot dm \int_0^\infty W(\xi) \cdot \cos m\xi \cdot d\xi, \quad \dots (28)$$

which reduces to

$$\psi_0 = \frac{4}{\pi} \int_0^\infty \frac{(mb - \lambda)s \sinh my - (s + mbc) \cosh my}{m^2[\lambda(sc + mb) + s^2 - m^2 b^2]} \times \sin mx \cdot dm \int_0^\infty W(\xi) \cos m\xi \cdot d\xi, \quad \dots (29)$$

in the case when f is large. Both (28) and (29) satisfy the necessary convergence conditions when $\lambda \neq 0$.

Formulæ for the Case when the Distribution of Load is Odd in x .

Let the load on $y=b$ be $W_1(x)$ per unit area, where $W_1(x)$ is an odd function such that $\int_0^\infty W_1(x) \cdot dx$ exists and is finite.

Instead of (13) we should now take $\chi = F(y) \sin mx$ and proceed as before, arriving finally at the desired stress function χ_1 , which would be given by

$$\chi_1 = \frac{2}{\pi} \int_0^\infty \frac{\lambda \sigma'(s + mbc) \cosh my + (\sigma mb - \lambda \sigma') msy \sinh my + \sigma(s + mbc) \sinh my - \sigma(s + mbc) my \cosh my}{m^2[\lambda \sigma'(sc + mb) + \sigma(s^2 - m^2 b^2)]} \times \sin mx \cdot dm \int_0^\infty W_1(\xi) \sin m\xi \cdot d\xi, \quad \dots (30)$$

while the corresponding displacement function, ψ_1 , would be

$$\psi_1 = \frac{-4}{\pi} \int_0^\infty \frac{(\sigma mb - \lambda \sigma') s \sinh my - \sigma(s + mbc) \cosh my}{m^2[\lambda \sigma'(sc + mb) + \sigma(s^2 - m^2 b^2)]} \times \cos mx \cdot dm \int_0^\infty W(\xi) \sin m\xi \cdot d\xi. \quad \dots (31)$$

The author has evaluated $(\bar{y}y)_{y=0}$ for a sufficient number of values of x' to indicate the general behaviour of the function. The results are

Fig. 1.

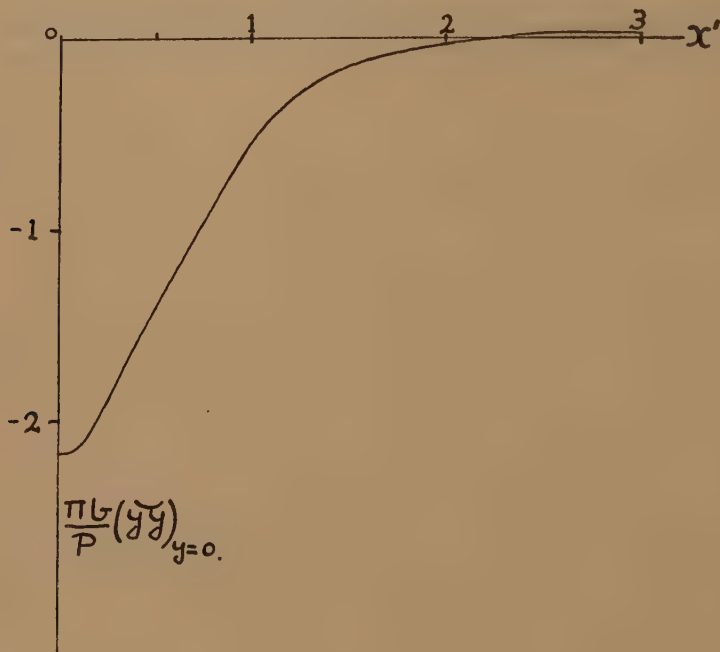
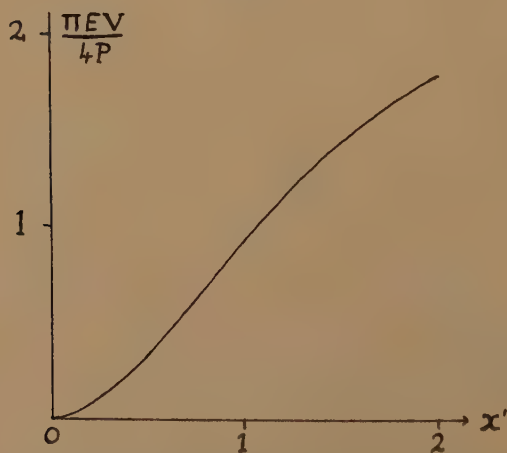


Fig. 2.



given in the table below and illustrated in the accompanying figures. In evaluating the integral (41), Filon's formula was used for the range

$t=0$ to $t=10.472$ (600°). After this stage the integrand may be taken as $(1+t)e^{-t} \cos tx'$ with an error of less than 1 in 10,000, and the integral from $t=10.472$ to infinity may therefore be evaluated by elementary means.

It will be noted that $(\bar{y}y)_{y=0}$ becomes positive for values of x' greater than about $2\frac{1}{4}$. This means that the assumption that the strip does not leave the foundation is incorrect. However, a rough quadrature shows that

$$\int_{-2.25b}^{2.25b} (\bar{y}y)_{y=0} \cdot dx = -1.01 \cdot P \text{ (Approx.)},$$

in other words the pressure between $x=\pm 2.25b$ is about 1 per cent. more than is necessary to balance the load. The tension over the interface (which is the tension that would have to be applied across the interface

Table showing the Deflexion at and the Stress across the Interface.

$x' \dots\dots\dots$	0	$\pm.5$	± 1	± 2	± 3
$\frac{\pi EV}{4P} \dots\dots\dots$	0	.3376	.9330	1.7947	—
$\frac{\pi b}{P} \cdot (\bar{y}y)_{y=0} \cdot$	-2.1709	-1.4095	-.5548	-.0248	.0244

in order to keep the strip in contact with the foundation) therefore amounts in the aggregate to only 1 per cent. (roughly) of the load, which is not large. The suggestion is, therefore, that in spite of this separation, which was not allowed for in the theory, the results of this section still give a good representation of the actual state of affairs.

We now turn to the deflexion of the beam along the interface. It will be readily found to be given by

$$\begin{aligned} E(\bar{v})_{y=0} &= -\frac{2P}{\pi} \int_0^\infty \frac{(s+mbc) \cdot \cos mx \cdot dm}{m[sc+mb+s^2-m^2b^2]} \\ &= -\frac{4P}{\pi} \int_0^\infty \frac{[(1+t)+(t-1)e^{-2t}]\cos tx'}{te^t[1-(1-2t+2t^2)e^{-2t}]} \cdot dt. \quad \dots \quad (42) \end{aligned}$$

This, as it stands, is a divergent integral, but it may be replaced by a convergent integral by finding, instead of $(\bar{v})_{y=0}$, the displacement, V , relative to the point of the interface immediately below the load.

We have

$$V = (\bar{v})_{y=0} - (\bar{v})_{x=y=0},$$

and therefore

$$EV = \frac{4P}{\pi} \int_0^\infty \frac{(1+t) + (t-1)e^{-2t}}{te^t[1 - (1-2t+2t^2)e^{-2t}]} (1 - \cos tx') \cdot dt, \quad . \quad . \quad . \quad (43)$$

which is convergent.

The Computation of EV.

We may write

$$\frac{1+t+(t-1)e^{-2t}}{te^t[1-(1-2t+2t^2)e^{-2t}]} = \phi(t) + \frac{1+t}{t} \cdot e^{-t}, \quad . \quad . \quad . \quad (44)$$

where

$$\phi(t) = \frac{1+t+(t-1)e^{-2t}}{te^t[1-(1-2t+2t^2)e^{-2t}]} - \frac{(1+t)}{t} e^{-t}. \quad . \quad . \quad . \quad (45)$$

The function $\phi(t)$ is regular at $t=0$ and we have, in fact,

$$\phi(0) = 0.$$

Substituting into (43) gives now

$$EV = \frac{4P}{\pi} \left\{ \int_0^\infty \phi(t) \cdot dt - \int_0^\infty \phi(t) \cdot \cos tx' \cdot dt + \int_0^\infty \frac{(1+t)e^{-t}}{t} \cdot (1 - \cos tx') \cdot dt \right\}. \quad . \quad . \quad (46)$$

Of the integrals occurring in (46), the first may be evaluated by any of the usual quadrature formulæ, the second by Filon's formula, and the third is elementary, its value being

$$\frac{1}{2} \log (1+x'^2) + 1 - \frac{1}{1+x'^2}.$$

We have, accordingly :

$$EV = \frac{4P}{\pi} \left\{ \int_0^\infty \phi(t) \cdot dt - \int_0^\infty \phi(t) \cdot \cos tx' \cdot dt + \frac{1}{2} \log (1+x'^2) + 1 - \frac{1}{1+x'^2} \right\}. \quad . \quad . \quad (47)$$

Values of V for several values of x' have been computed and are given in the table, besides being illustrated graphically. Fig. 3 illustrates the relation between $(\widetilde{yy})_{y=0}$ and V and is, in the author's opinion, of particular interest in that it shows that the assumption of the elementary theory that the relationship between $(\widetilde{yy})_{y=0}$ and V is linear is very far from being true if the depth of the foundation is large.

The Case when f is small.

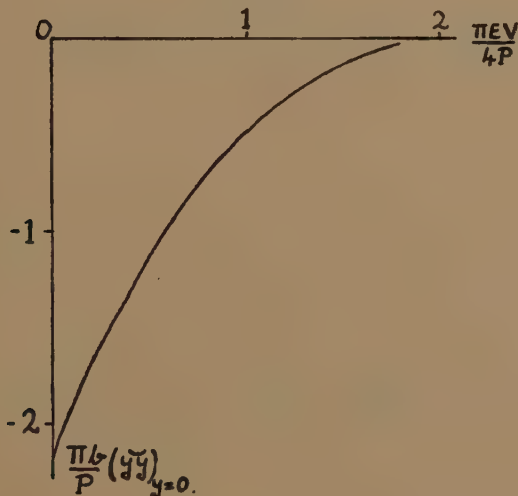
Let us return now to equation (12). If $(\bar{y}y)_{y=0}$, $(\bar{v})_{y=0}$, and their differential coefficients are bounded, we may expand the operators in ascending powers of f . If this is done it will be found that, on assuming f to be small and retaining only the leading terms, equation (12) reduces to

$$fD^2(\bar{y}y)_{y=0} = D^2E'(\bar{v})_{y=0}. \quad (48)$$

If, now, we use the notation of the elementary theory (*i. e.*, denote the deflexion of the strip by y , etc.) this takes the form

$$fD^2 \cdot \frac{EI}{c'} \cdot D^4y + D^2E'y = 0,$$

Fig. 3.



and this may be written

$$\left(D^4 + \frac{c'E'}{EI f}\right) D^2y = 0. \quad (49)$$

The solution of this equation is, apart from a linear term, exactly that of Mr. Coe's paper, and indicates that the constant k used in that paper is simply $\frac{c'E'}{f}$, which is what would be anticipated from simple physical considerations.

The stress and displacement functions appropriate to the case when f is of the order of smallness assumed here are readily obtained from (25) and (28) by putting $\frac{\sigma}{\sigma'} = \frac{mf}{2}$, a fact readily verified by comparing (48)

with (19). It is then easy to show that the relationship between $(\bar{v})_{y=0}$ and $(\bar{y}y)_{y=0}$ is, apart from a trivial term, linear. The general conclusion is, therefore, that the elementary theory gives, within its usual limitations, a reasonably good approximation to the truth if the thickness of the foundation is small. On the other hand, as has been seen, if the thickness of the foundation is large, the assumption of linearity between pressure and deflexion at the interface is in considerable error.

Note.—The terms “large” and “small” as applied to f in this paper have been used rather loosely and, in a sense, involve a knowledge of the results which would be obtained by using the general forms of the stress and displacement functions (25) and (28). Thus f is “small” if (48) gives a sufficiently close approximation to the results which would be got from (25) and (28), and is “large” if the same is true of (25*a*). Considerations of brevity preclude a more detailed consideration of this point, but it will actually be found safe to interpret these terms as meaning “large” or “small” compared with b .

References.

- (1) R. Thornton Coe, “The Bending of Beams on Elastic Foundations,” *Phil. Mag.* xxv. no. 166 (Jan. 1938).
- (2) See A. E. H. Love, ‘*Mathematical Theory of Elasticity*,’ 4th ed., Cambridge, 1927, p. 138.
- (3) For details of a similar phenomenon in the case of an inelastic foundation see L. N. G. Filon, *Phil. Trans. A*, cci. (1903).
- (4) For further information as to stress functions of this type see W. M. Shepherd, *Journ. Lond. Math. Soc.* iv. pt. 3, pp. 213–19.
- (5) L. N. G. Filon, “On a Quadrature Formula for Trigonometrical Integrals,” *Proc. Roy. Soc. Edin.* xlix. pp. 38–47 (1929).

Mathematical Department,
Technical College, Brighton.

IV. *A further Note on the Definition and Determination of Mass in Newtonian Mechanics.*

By C. G. PENDSE, M.A., Ph.D.*

[Received August 23, 1938.]

IN a former paper † I discussed the problem of the definition and determination of mass in Newtonian mechanics, and examined Mach's definition of the ratio of the masses of two bodies (particles), namely, the negative inverse ratio of their mutually induced accelerations. It was shown that though Mach's definition, in its original form, was applicable to the case of two bodies (particles) forming an "isolated" system, it was possible to argue with the help of the laws of motion and the corollary to the second law that the ratios of the masses of pairs of particles in a system composed of any number of particles were *conceptually* perfectly definite. It was further concluded that, nevertheless, Mach's definition did not enable an observer (a "Newtonian" observer) to determine the ratios of the masses of pairs of particles, or the forces between them, in an "isolated" system when it was composed of (i.) more than four particles, the accelerations at one instant alone being considered, and (ii.) more than seven particles, the number of instants at which the accelerations are taken being arbitrary.

In the present note, which consists of Part I. (§ 1) and Part II. (§§ 2-4), the question is considered further.

In Part I. (§ 1) below it is shown that the ratio of the masses of two particles, as defined by Mach, does not have the *same* constant value for *all* observers.

In § 4 of the former paper it was assumed that the observer of the motion of the particles could regard it as an "isolated" system. In Part II. below the two points of view from which an observer can look at the problem are discussed and considered separately.

In § 1 of I. the well-known portions of mechanics were briefly discussed; in § 2 Mach's definition was stated, and in § 3 it was examined when the system contained more than two particles. That portion is not repeated here.

* Communicated by Prof. W. M. Smart.

† "A Note on the Definition and Determination of Mass in Newtonian Mechanics," *Phil. Mag.* (7) xxiv. pp. 1012-1022 (1937).

This is referred to as either "the former paper" or I.

PART I.

§ 1. *The Variability of the Constant Ratio of the Masses of Two Particles, as defined by Mach, with Observers.*

As pointed out in §§ 2 and 3 of I., Mach had assumed that the two particles formed an "isolated" system in so far as the frame of reference was concerned, and he defined the ratio of their masses as the "negative inverse ratio of their mutually induced accelerations," which he postulated to be constant in time and positive. We have to examine if, when there exists one such frame, there also exist other frames such that in the case of each the system of the two particles can be deemed to be "isolated" and the negative inverse ratio of the accelerations would be constant in time and positive, and if it be so, whether the constants for the various frames would differ from or equal one another.

Let P_1 and P_2 be two particles and $Oxyz$ a frame of reference such that their accelerations relative to it are always parallel to the straight line joining them. Let f_1 and f_2 be the accelerations of P_1 and P_2 parallel to $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_2P_1}$ respectively; and let $\frac{f_2}{f_1}$ be constant in time and positive and equal to m_{12} .

Let $O'x'y'z'$ be another frame of reference moving relative to $Oxyz$ with acceleration but without rotation. At each instant let the acceleration of O' relative to O be f parallel to $\overrightarrow{P_1P_2}$, f being positive or negative. Then the accelerations of P_1 and P_2 relative to $O'x'y'z'$ will also be parallel to $\overrightarrow{P_1P_2}$; let these be f_1' and f_2' , parallel to $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_2P_1}$ respectively. Then we have

$$f_1' = f_1 - f,$$

$$f_2' = f_2 + f,$$

and

$$\frac{f_2'}{f_1'} = \frac{f_2 + f}{f_1 - f}.$$

Then $\frac{f_2'}{f_1'}$ will also remain constant in time and be equal to an arbitrary positive number m'_{12} , if

$$\frac{f_2 + f}{f_1 - f} = m'_{12},$$

i. e., if

$$\frac{f_2}{f} \left(\frac{1 + \frac{f}{f_2}}{1 - \frac{f}{f_1}} \right) = m'_{12},$$

e., if
$$m_{12} \left(\frac{1 + \frac{f}{f_2}}{1 - \frac{f}{f_1}} \right) = m'_{12},$$

i. e., if
$$f = \frac{m'_{12} - m_{12}}{\frac{m_{12}}{f_2} + \frac{m'_{12}}{f_1}}. \quad \dots \dots \dots (A)$$

Thus we see that if the accelerations of P_1 and P_2 relative to O are always parallel to the line joining them, and if their negative inverse ratio remains constant in time and positive, their accelerations relative to O' will also be always parallel to the line joining them, and their negative inverse ratio will remain constant in time and equal an arbitrarily chosen positive number provided the acceleration of O' relative to O be suitably chosen *.

* The argument may be made clearer as follows :—

Let the accelerations of P_1 and P_2 relative to O be f_1 and f_2 parallel to $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_2P_1}$ respectively, and let the acceleration of O' relative to O be f parallel to $\overrightarrow{P_1P_2}$, where

$$f = \frac{m'_{12} - m_{12}}{\frac{m_{12}}{f_2} + \frac{m'_{12}}{f_1}},$$

m_{12} and m'_{12} being some positive constants. The accelerations of P_1 and P_2 relative to O' will then be f'_1 and f'_2 parallel to $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_2P_1}$ respectively, where $f'_1 = f_1 - f$ and $f'_2 = f_2 + f$. Then

$$\begin{aligned} \frac{f'_2}{f'_1} &= \frac{f_2 + f}{f_1 - f}, \\ &= \frac{f_2 + \frac{m'_{12} - m_{12}}{\frac{m_{12}}{f_2} + \frac{m'_{12}}{f_1}}}{f_1 - \frac{m'_{12} - m_{12}}{\frac{m_{12}}{f_2} + \frac{m'_{12}}{f_1}}} \\ &= \frac{m'_{12}}{m_{12}} \times \frac{f_2}{f_1}, \text{ on simplification.} \end{aligned}$$

Hence, if $\frac{f_2}{f_1}$ remains constant in time and equals m_{12} , $\frac{f'_2}{f'_1}$ also remains constant in time and equals m'_{12} . (m_{12} and m'_{12} are arbitrary positive constants, and $\frac{f_2}{f_1}$ and $\frac{f'_2}{f'_1}$ are the negative inverse ratios of the accelerations of P_1 and P_2 relative to O and O' respectively.)

Suppose now that a frame $Oxyz$ is such that the accelerations of P_1 and P_2 are at every instant parallel to the line P_1P_2 and that their negative inverse ratio remains constant in time and equals a positive number m_{12} ; and we consider a frame $O'x'y'z'$ of the type mentioned above, its acceleration relative to $Oxyz$ being such that the accelerations of P_1 and P_2 relative to O' are at every instant parallel to the line P_1P_2 and that their negative inverse ratio remains constant in time and equals an arbitrarily chosen positive number m_{12}' . We further consider observers fixed in the frames $Oxyz$ and $O'x'y'z'$, called for the sake of brevity observer O and observer O' . Each observer can only observe the motions. He cannot observe the ratio of the masses of the particles or the force between them: he can only aspire to infer them.

As each observer finds that the accelerations of P_1 and P_2 are at every instant parallel to the line P_1P_2 and are oppositely directed and that their negative inverse ratio remains constant and positive, he will be unable to say that there are "external" forces on the system; he will be unable to differentiate, in the case of each particle, between the acceleration produced by the action of the other particle and that due to an "external" force. Each will be obliged to conclude that the system is "isolated" and that the positive constant negative ratio determined by him is the ratio of the masses of the particles. Thus each observer will have his own value of the ratio. The forces for the two observers can then be determined easily.

From the relation (A) it is seen that when $f=f_1$, $m'_{12}=\infty$, and when $f=-f_2$, $m'_{12}=0$. The two limiting cases mean that in the first case the observer is at rest relative to P_1 , and that in the second he is at rest relative to P_2 . Consequently in each of the two cases one of the particles is at rest, and therefore the question of determining the ratio of their masses from the accelerations of both does not arise.

Hence we conclude that when there is one observer who can say that a system of two particles is "isolated" and find the ratio of their masses, there exists an infinity of observers each of whom can regard the system as "isolated" and have his own determination of the ratio of their masses, as defined by Mach, the domain of the values of the ratio being the set of positive numbers*.

* In §3 of I. the question how far Mach's definition was applicable when the system contained more than two particles was considered. In view of the conclusion reached here we can say: In the case of a system composed of more than two particles, whose motion is referred to a Newtonian frame, it is possible to argue that if the ratio of the masses of two particles be defined as the "negative inverse ratio of their mutually induced accelerations" the ratios of the masses of the pairs of particles in the system are *conceptually* perfectly definite *relative* to the frame of reference.

PART II.

§2. *The Two Aspects of the Problem of the Determination of the Ratios of the Masses of the Particles of a System by an Observer.*

Using the suffixes $\lambda, \mu, \nu = 1, 2, \dots, n$, we consider a system of n particles P_λ . The motions of the particles are referred to and observed by an observer fixed in a frame $Oxyz$. He is supposed to know the motion of every particle in the system.

Let $\vec{e}_{\lambda\mu}$ be the unit vector in the direction $P_\lambda \rightarrow P_\mu$ and let \vec{f}_λ be the acceleration of the particle P_λ at any instant. That part of \vec{f}_λ which is "due to the action of the particle P_μ on P_λ " is along the line $P_\lambda P_\mu$; we denote it by $\phi_{\lambda\mu} \vec{e}_{\lambda\mu}$.

The questions which arise in connexion with the relation of the observer and the frame to the system of particles are:—

1. Is the observer to assume that the system can be considered to be "isolated" in so far as the frame of reference is concerned?
2. Is he to assume the property of inertia, *i. e.*, the existence of n positive numbers, each number being associated with one particle, one of the numbers being chosen arbitrarily?
3. Is he to assume the laws of motion?

If he makes these assumptions, his work reduces to that of computing the ratios of the masses of the particles and the interactions between them: his problem is one of computation. If, however, he does *not* make these assumptions, his work is to discover the existence of the constant ratios of masses and to determine whether the system is "isolated" or not; when he has found out the existence of these ratios and that the system can be deemed to be "isolated," he has to determine the ratios of the masses of the particles and the forces between them: in this case his problem is heuristic rather than computational.

We consider, therefore, the two aspects, namely, (i.) the heuristic, and (ii.) the computational.

§ 3. *The Heuristic Aspect.*

Since the observer can only observe the motions he can only have Mach's definition of the ratio of the masses of two particles and the hypotheses accompanying it to decide, at least provisionally, whether or not the system can be deemed to be "isolated" and he is a Newtonian observer, and to determine the ratios of the masses of the particles and the forces between them.

The operations which he has to perform are as follows :—

(1) He has to find out if he can uniquely express the acceleration of every particle as the sum of vectors parallel to the straight lines joining that particle to others at every instant.

(2) Having carried out the first operation and seen that the unique expression is possible he has to find out if in the case of every pair the ratio

$$\frac{\Phi_{\mu\lambda}}{\Phi_{\lambda\mu}}, \text{ where } \vec{f}_\lambda = \sum_{\substack{\mu=1 \\ \mu \neq \lambda}}^n \Phi_{\lambda\mu} \vec{e}_{\lambda\mu},$$

remains constant in time and is positive.

(3) Assuming that he has carried out the operations in (1) and (2) and found that the ratios are constant, let $m_{\lambda\mu} = \frac{\Phi_{\mu\lambda}}{\Phi_{\lambda\mu}}$, which is constant in time and positive. He has then to verify that

$$m_{\lambda\nu} = m_{\lambda\mu} \times m_{\mu\nu}. \quad \dots \dots \dots (B)$$

If he can carry out the operations described in (1), (2), and (3) above, he can identify $\Phi_{\lambda\mu}$ with $\phi_{\lambda\mu}$ and assume that the system is “isolated”; he would then have no means of discriminating between the accelerations due to the “external” forces and those due to the “internal” ones, and he would be able to assume that the frame of reference is Newtonian, the system being “isolated.”

We assume that he can carry out the operations. He knows the ratios of the masses of the pairs of particles; then, by assigning an arbitrary positive value to the mass of one of the particles, he can specify those of others and also determine the forces between the pairs.

If he can carry out the operations for a large number of instants he can assume, at least provisionally, that the system is “isolated” and the frame is Newtonian, and determine the masses of the particles and the forces at the corresponding instants between them.

It is to be noticed that the relation (B) is an independent axiom and that it is not contained implicitly or explicitly in Mach’s definition of the ratio of masses. When the observer finds it possible to determine the constants $m_{\lambda\mu}$ they are determined from the observed accelerations. Consequently there is no *a priori* reason why the relation (B) should be satisfied automatically. However, the reason for the necessity of the axiom (B) is as follows.

We choose a particle P_1 and assign to it an arbitrarily chosen positive number m_1 , which we call its mass. The masses of the other particles, m_2, \dots, m_n , are given by the relations

$$m_2 = m_1 m_{21}, \quad m_3 = m_1 m_{31}, \quad \dots, \quad m_n = m_1 m_{n1}.$$

Thus we obtain the masses m_1, m_2, \dots, m_n , by selecting m_1 arbitrarily and using the ratios m_{21}, \dots, m_{n1} . But as m_1, \dots, m_n are algebraic numbers, they satisfy the general relation

$$\frac{m_\lambda}{m_\nu} = \frac{m_\lambda}{m_\mu} \times \frac{m_\mu}{m_\nu}.$$

Hence, if in the general case $\frac{m_\lambda}{m_\mu}$ is the same as the ratio $m_{\lambda\mu}$ as determined by observation, while the m_λ 's are determined from m_{21}, \dots, m_{n1} , we must have the relation (B), namely,

$$m_{\lambda\nu} = m_{\lambda\mu} \times m_{\mu\nu}.$$

The heuristic aspect corresponds to case (I.) of § 4 on pp. 1017-1019 of the former paper. Considering the results proved there the observer will never be able to carry out the operations described above in (1), (2), and (3), when $n > 4$ and even when $n = \begin{Bmatrix} 3 \\ 4 \end{Bmatrix}$ if the particles always remain $\begin{Bmatrix} \text{collinear} \\ \text{coplanar} \end{Bmatrix}$. In these cases he will be unable to find the ratios $\frac{\Phi_{\mu\lambda}}{\Phi_{\lambda\mu}}$ at any instant.

Hence, when $n > 4^*$, the observer would never be able to decide whether or not the system is "isolated" and to determine the ratios of the masses of the particles or the forces between them. When $n = \begin{Bmatrix} 3 \\ 4 \end{Bmatrix}$ he would be able to find out whether or not the system is "isolated" only if the particles do not always remain $\begin{Bmatrix} \text{collinear} \\ \text{coplanar} \end{Bmatrix}$, and if the system be found to be "isolated" he would be able to determine the ratios of the masses and also the interactions at instants when they are not $\begin{Bmatrix} \text{collinear} \\ \text{coplanar} \end{Bmatrix}$. When $n=2$ he would be able immediately to decide whether or not the system is "isolated"; and if it be "isolated" he would also find the ratio of their masses and the interaction between them †.

* When the observer is unable to carry out the operation described in (1) above for each pair of particles there will be an infinity of numbers similar to $\Phi_{\mu\lambda}/\Phi_{\lambda\mu}$ at every instant. And it will be impossible for him to find out unique ratios which remain constant for observations at a number of instants.

† In the first place, when $n=2$, the accelerations of the particles must always be parallel to the straight line joining them and in opposite directions if the system is "isolated." Secondly the negative inverse ratio of the accelerations must be constant in time and positive.

§4. The Computational Aspect.

In this case the observer assumes that the system of particles is an "isolated" one in so far as the frame of reference is concerned; he assumes the laws of motion and the existence of constant mass coefficients. He only wants to compute the ratios of the masses of the particles and the forces between them. He would be able to use the equations of motion.

This aspect corresponds to case (II.) in §4 on p. 1019 of the former paper.

Let m_λ be the mass of the particle P_λ in terms of some arbitrary unit, and let $\vec{F}_{\lambda\mu} e_{\lambda\mu}$ be the force on P_λ due to P_μ . By the third law of motion $\vec{F}_{\lambda\mu} = \vec{F}_{\mu\lambda}$.

At each instant there will be n vector equations of motion

$$m_\lambda \vec{f}_\lambda = \sum_{\substack{\mu=1 \\ \mu \neq \lambda}}^n \vec{F}_{\lambda\mu} e_{\lambda\mu} \dots \dots \dots (4.1/\lambda)$$

At each instant there will be not more than $3n$ distinct homogeneous linearly independent algebraic equations, the number of unknowns being $n + \frac{1}{2}n(n-1)$, the masses m_λ , and the interactions $\vec{F}_{\lambda\mu}$; they will have an indeterminate solution when $n > 5$.

For different instants the $\vec{F}_{\lambda\mu}$'s will be different, but the m_λ 's will remain the same. Hence, corresponding to r instants, there will be not more than $3nr$ distinct homogeneous linearly independent algebraic equations, while the number of unknowns will be $n + \frac{1}{2}n(n-1)r$; so far as *all* the unknowns* are concerned the equations will have an indeterminate solution whatever the value of r may be when $n > 7$, since $n - 1 + \frac{1}{2}n(n-1)r > 3nr$ for all values of r when $n > 7$. It is not possible, however, to conclude immediately that the ratios of the masses will be indeterminate when $n > 7$; for, from the equations (4.1/ λ) we derive, for each instant, the vector equations

$$\sum_{\lambda=1}^n m_\lambda \vec{f}_\lambda = 0, \dots \dots \dots (4.2)$$

and

$$\sum_{\lambda=1}^n m_\lambda [\vec{OP}_\lambda \times \vec{f}_\lambda] = 0. \dots \dots \dots (4.3)$$

* We have taken homogeneous equations for the sake of simplicity. If we consider only the ratios of the masses, there will be $n - 1 + \frac{1}{2}n(n-1)r$ unknowns and not more than $3nr$ distinct non-homogeneous linear algebraic equations, as in case (II.) in §4 of the former paper. This, however, makes no difference to the conclusion regarding the nature of the solution so far as all the variables are concerned.

The equations (4.2) and (4.3)* correspond to the conservation of linear momentum and angular momentum respectively, and do not contain the $F_{\lambda\mu}$'s. If we consider $\left\{ \begin{smallmatrix} \text{one} \\ \text{both} \end{smallmatrix} \right\}$ of the vector equations (4.2) and (4.3) we have not more than $\left\{ \begin{smallmatrix} \text{three} \\ \text{six} \end{smallmatrix} \right\}$ distinct homogeneous linearly independent algebraic equations at each instant, and accordingly we would have not more than $\left\{ \begin{smallmatrix} 3r \\ 6r \end{smallmatrix} \right\}$ distinct homogeneous linearly independent algebraic equations for r instants, containing the m_λ 's alone; and assuming that the equations for r instants are linearly independent, these would have a determinate solution for the ratios of the m_λ 's if

$$\left\{ \begin{smallmatrix} 3r > n-1 \\ 6r > n-1 \end{smallmatrix} \right\}.$$

Thus it *might* be possible for the observer to determine the ratios of the masses when $n \geq 7$, provided r is sufficiently large; the forces between the particles would, however, remain indeterminate †.

It does not, however, follow that the ratios of the masses can *always* be determined when accelerations at a sufficiently large number of instants are taken into account. A simple example will illustrate the position.

Let the particles‡ occupy the vertices of a regular polygon of n sides inscribed in a circle of fixed radius a and with centre fixed in the frame of reference and revolving with uniform angular velocity ω ; let C be the centre of the circle. The acceleration of the particle P_λ is $a\omega^2$ in the

* I am grateful to Prof. E. A. Milne for drawing my attention, after the publication of the former paper, to the importance of the fact that the equation (4.2) above is an eliminant of the equations (4.1/ λ) and is free from the $F_{\lambda\mu}$'s. For particular values of n it is possible to find, using equations (4.2) alone, the minimum values of r necessary for determining the ratios of the masses in appropriate cases.

† This conclusion is rather different from that arrived at in case (II.) of §4 of the former paper. It was stated there that the observer would never be able to determine the ratios of the masses of the particles or the forces between them when $n > 7$. The new conclusion reached here is that he *might* be able, in some cases, to determine the ratios of masses by taking into account accelerations at a sufficiently large number of instants, but that the forces would remain indeterminate when $n > 7$.

‡ This state of motion is actually possible when the masses of the particles are equal and the particles attract each other according to the law of gravitation. This is a model of a ring of Saturn without the planet at the centre. Even if we were to assume that there is a particle of mass M at the centre C the analysis for the determination of the ratios of the masses from observed accelerations, given below, remains intact.

(2) Confining attention to one frame of reference and an observer fixed in it, the problem of the determination of the ratios of the masses of a system of n particles is examined. There are two aspects of the problem : (i.) heuristic, and (ii.) computational.

(i.) The observer has to find out, from his knowledge of the motions, whether or not the system can be deemed to be "isolated" ; he has to establish the existence of the constant ratios of masses and to ascertain if he can use the laws of motion. He cannot carry out the necessary operations at all, and consequently cannot determine the ratios of the masses of the particles or the forces between them when $n > 4$, and even when $n = \begin{Bmatrix} 3 \\ 4 \end{Bmatrix}$, if the particles always remain $\begin{Bmatrix} \text{collinear} \\ \text{coplanar} \end{Bmatrix}$.

(ii.) The observer assumes that he can regard the system as "isolated," that there exist constant mass coefficients, one associated with each particle and one of them being chosen arbitrarily, and that he can use the laws of motion. When $n > 7$ he would be unable to determine uniquely the forces between the particles, though he *might* be able, in some cases, to determine the ratios of the masses by considering accelerations at a sufficiently large number of instants.

Downing College,
Cambridge.

V. *Three Elementary Examples of the Uncertainty Principle.*

By D. S. KOTHARI, Ph.D. (Cambridge)*.

[Received July 11, 1938.]

THE uncertainty principle states the inherent (ultimate) limitations on the possibilities of observation of the states of atomic particles, and forms an essential feature of modern quantum theory. A number of illustrative examples of the principle are to be found in current literature⁽¹⁾, and the present note deals with three elementary examples which, it is believed, have not been discussed (at least explicitly) previously. The first example deals with the "microscope," but from the point of view of its "depth of focus" and not, as is usual, its resolving power. The second example is concerned with the measurement of angle, and the third with the determination of velocity by a direct method—in principle the same as recent methods for the determination of the velocity of atomic beams.

1. *Microscope and the Depth of Focus.*—It is easily seen that the "depth of focus" of a microscope is given by

$$\epsilon = \frac{\lambda}{(1 - \cos \alpha)},$$

where α is the semi-angular aperture of the objective. Let AB in fig. 1 (a) be the objective of the microscope and suppose that a point object at O_1 produces a maximum at I, and a neighbouring point object O_2 gives rise to a minimum at the same point I. Then O_1O_2 is the depth of focus. As O_1 produces a maximum at I the optical paths O_1AI and O_1BI are equal, and as O_2 gives rise to a minimum at I, the optical paths O_2AI and O_2BI differ by λ , λ being the wave-length of light †. We thus have

$$O_2AI - O_2BI = \lambda, \quad \dots \dots \dots (1)$$

and as

$$O_1AI = O_1BI,$$

$$(O_2AI - O_1AI) - (O_2BI - O_1BI) = \lambda,$$

or

$$\epsilon \equiv O_1O_2 = \frac{\lambda}{(1 - \cos \alpha)} \dots \dots \dots (2)$$

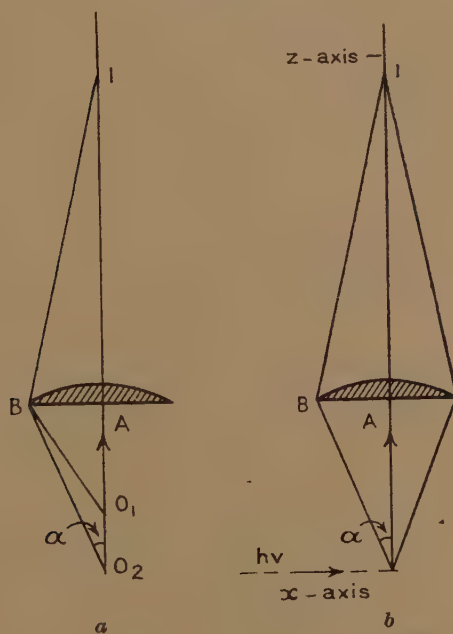
* Communicated by the Author.

† For minima at I the path difference required between O_2AI and O_2BI is λ and not $\lambda/2$. If we suppose the semi-aperture AB to be divided into two halves, then every ray in the first half will differ from a corresponding ray in the lower half by a path difference of $\lambda/2$, and their effect at I will thus cancel out.

Fig. 1 (b) illustrates the microscope being used to observe a particle, say, an electron. We take the incident photon along the x -axis, and after being scattered by the electron it enters the microscope. The scattered photon will make with the Z -axis (the axis of the microscope) any angle between zero and α , its exact value being uncertain, and thus the z -component of the (recoil) momentum imparted to the electron by the photon will lie anywhere between

$$\frac{h\nu}{c} \quad \text{and} \quad \frac{h\nu}{c} \cos \alpha.$$

Fig. 1.



Thus the uncertainty in the z -component of the momentum will be given by

$$\Delta p_z \gtrsim \frac{h\nu}{c} (1 - \cos \alpha). \quad (3)$$

The uncertainty Δz in the distance of the electron from the microscope, i. e., the z -coordinate of the electron, will be $\Delta z \gtrsim \epsilon$, hence we have

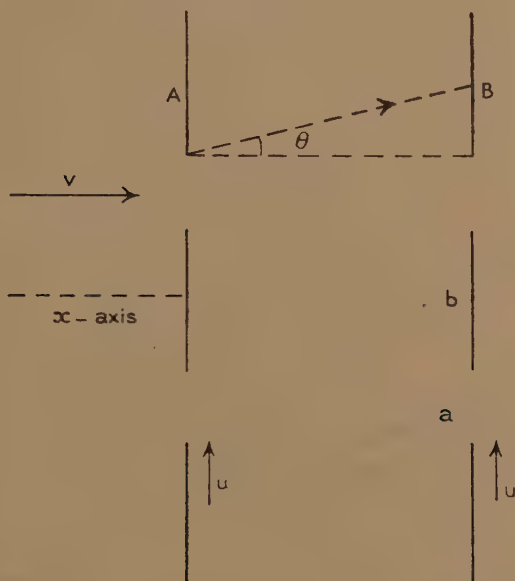
$$\Delta p_z \Delta z \gtrsim \frac{h\nu}{c} (1 - \cos \alpha) \frac{\lambda}{1 - \cos \alpha} = h, \quad (4)$$

which is the uncertainty relation.

of the method is illustrated in fig. 3. A and B represent two parallel rows of slits which can be moved along their own lengths with equal velocities. The distance between them is l . The width of the slit is a , and b is the width of the opaque space between two adjacent slits. The slits in A are exactly opposite the slits in B (and similarly for the opaque spaces). In the sequel A, B will be sometimes referred to as gratings.

Let a particle, say an electron, moving parallel to the x -axis (*i. e.*, normally to A, B), enter one of the slits in A. Let u be the (minimum)* velocity of A, B, such that the electron is not stopped by an opaque

Fig. 3.



space but passes through a slit in grating B. Under such conditions the velocity v of the electron will be related with u by the relation

$$\frac{l}{v} = \frac{b}{u} \quad \dots \dots \dots (8)$$

A little consideration, however, shows that the distance actually moved by the grating B (and so also the grating A) while the electron travels the distance l will not be exactly b but vary within certain limits $b \pm \epsilon$, due to two reasons: (1) the width a of the slit, and (2) the diffraction of the electron-wave as it passes through the slit. The angle of diffraction

* We shall assume for simplicity that the slit in B through which the electron passes is adjacent to the slit opposite to the slit in A by which the electron entered.

θ will be approximately equal to λ/a , λ being the wave-length of the electron-wave. The effect of (1) will be to make $\epsilon \gtrsim a$, and the effect of (2) will be to make $\epsilon \gtrsim l\theta$, and hence

$$\epsilon^2 \gtrsim al\theta = l\lambda. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

This uncertainty of 2ϵ in the distance moved by the grating B, during the time the electron travels the distance between the gratings, introduces in the velocity measurement an error Δv , which because of (8) will be

$$\frac{\Delta v}{v} = \frac{2\epsilon}{b}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

This uncertainty Δv in the velocity will render any previous knowledge of the x -coordinate of the electron uncertain by at least the amount $\Delta v \cdot \frac{l}{v}$, v being the time the electron spends in passing through the apparatus, *i. e.*,

$$\Delta x \gtrsim \Delta v \cdot \frac{l}{v}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

and hence using (8) and (10) we have

$$\Delta p_x \Delta x \gtrsim m \frac{l}{v} (\Delta v)^2 = 4h \left(\frac{v}{u} \right)^2 \frac{\epsilon^2}{\lambda l}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

since $\lambda = h/mv$.

One further point has to be considered in this connexion. As the electron passes through the slit in grating B (and also the slit in A), its wave-train after it has emerged from the apparatus cannot exceed in length $\frac{a\theta}{u}$, $\frac{a}{u}$ being the time for which the wave-train moving along the x -axis is allowed to pass through the slit. This implies an uncertainty Δv_0 in the velocity such that *

$$\Delta v_0 \gtrsim \frac{h}{m \left(v \frac{a}{u} \right)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

* If S is the length of the wave-train, then the wave-length will lie in the interval $\Delta\lambda$ such that

$$\frac{S}{\lambda} - \frac{S}{\lambda + \Delta\lambda} \sim 1, \text{ or } \frac{\Delta\lambda}{\lambda^2} S \sim 1.$$

As $\lambda = h/mv$, we have $\frac{\Delta\lambda}{\lambda} = \frac{\Delta v}{v}$, and substituting for $\Delta\lambda$ in the above we have

$$S \sim \frac{h}{m \Delta v}.$$

For the experiment above discussed to have any meaning this uncertainty Δv_0 (which enters at the end of the experiment) must not exceed Δv .

Writing $k = \frac{\Delta v_0}{\Delta v}$, and using (10), (13), and (8), we obtain

$$k = \frac{\Delta v_0}{\Delta v} = \frac{l\lambda}{2\epsilon a} \left(\frac{u}{v}\right)^2 \quad (14)$$

We substitute for $\left(\frac{v}{u}\right)$ in (12) from (14) and have

$$\Delta p_x \Delta x \gtrsim 2\hbar \frac{\epsilon}{a} \frac{1}{k} \quad (15)$$

Remembering that $\epsilon \gtrsim a$ and $k \gtrsim 1$, (15) shows that

$$\Delta p_x \Delta x \gtrsim 2\hbar, \quad (16)$$

which is the uncertainty relation*.

* In case $\Delta v_0 \gtrsim \Delta v$, the uncertainty in the x -coordinate will be $\frac{l}{v} \Delta v_0$ (and not $\frac{l}{v} \Delta v$), and we shall have

$$\Delta p_x \Delta x \gtrsim m \frac{l}{v} (\Delta v_0)^2 = \hbar \frac{l\lambda}{a^2} \left(\frac{u}{v}\right)^2.$$

Substituting for $\left(\frac{u}{v}\right)$ from (14), we obtain

$$\Delta p_x \Delta x \gtrsim k \frac{2\epsilon}{a} \hbar,$$

and as $\epsilon \gtrsim a$, and for this case $k \sim 1$, we have

$$\Delta p_x \Delta x \gtrsim 2\hbar.$$

Reference.

- (1) Heisenberg, 'The Physical Principles of the Quantum Theory,' Chapter II. (1930). Darwin, Proc. Roy. Soc. A, cxxx. p. 632 (1930). Also see Williams, 'Science Progress,' xxxi. p. 14 (1936).

Physics Department,
University of Delhi.

VI. *A Hot Wire Method for the Thermal Conductivities of Gases.*

By G. G. SHERRATT, B.A., and EZER GRIFFITHS, D.Sc., F.R.S., Physics Department, National Physical Laboratory, Teddington, Middlesex *.

[Received October 15, 1938.]

FOR the determination of the thermal conductivity of gases there are two methods in common use. One utilizes a heater in the form of a flat plate whilst in the other the heater takes the form of a wire. When conductivity data are required over a wide range of temperature the hot wire method is the more adaptable. There are, however, certain disadvantages in this method: the correction for the temperature discontinuity at the gas-metal surface becomes uncertain when the wire is fine, and the alternative, the use of thick wires, has hitherto involved corrections of considerable magnitude, for the heat conducted to the cold ends. In the method now to be described a thick wire is used, and heat conduction to the ends eliminated by auxiliary sources of heat:

The wire is provided with two potential points, so that the electrical energy dissipated in a specified length of wire can be determined in the usual manner. Uniformity of temperature over this length of wire is obtained by supplying heat at a point near each end by means of auxiliary heaters, consisting of two loops welded at their mid-points to the wire. The loops are heated by independent sources of electrical supply. The first application of the apparatus has been to the determination of the thermal conductivity of carbon dioxide, dry air (free from carbon dioxide) and dichloro-difluoro-methane. In recent years the latter gas has been used to a considerable extent in refrigeration work for "air conditioning" in place of ammonia and carbon dioxide in the compression type of machine.

Commercially the gas is known as Freon or F 12.

Description of Apparatus.

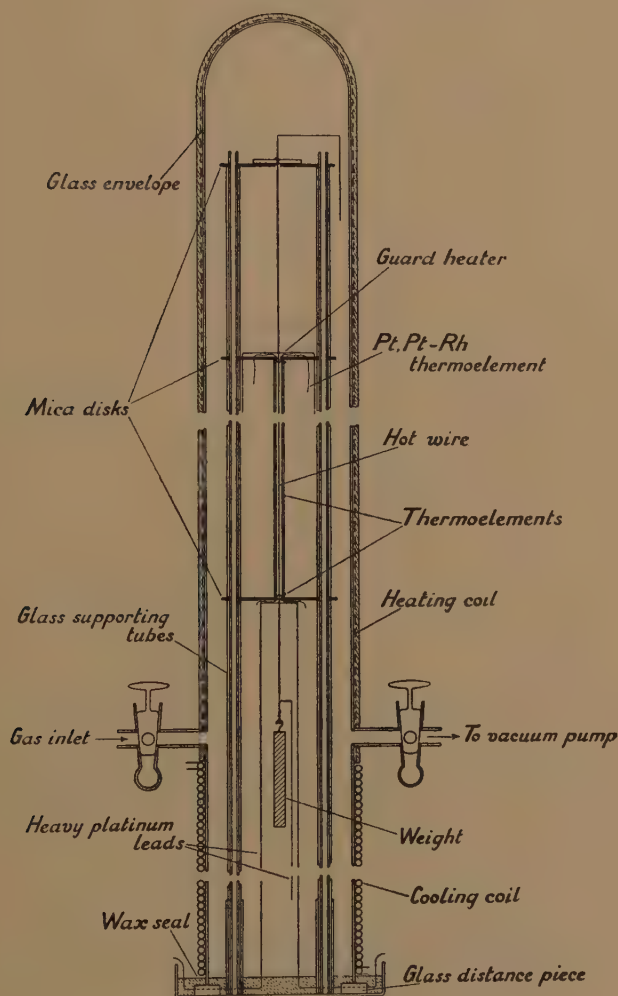
The apparatus is constructed mainly of Pyrex glass. It consists in its essentials of a supporting framework of glass tubing, which is mounted vertically in a flat-bottomed dish. The framework carries three mica disks, which are so arranged that their centres are in a vertical line and their planes horizontal.

From the centre of the upper disk the "hot wire" hangs vertically and is kept taut by a weight. It passes through closely fitting holes

* Communicated by the Authors.

in the centre of the two lower disks, and a glass tube which surrounds the hot wire is held between the disks. Friction between the ends of the tube and the disks serves to hold the former in position with respect to the wire.

Fig. 1.



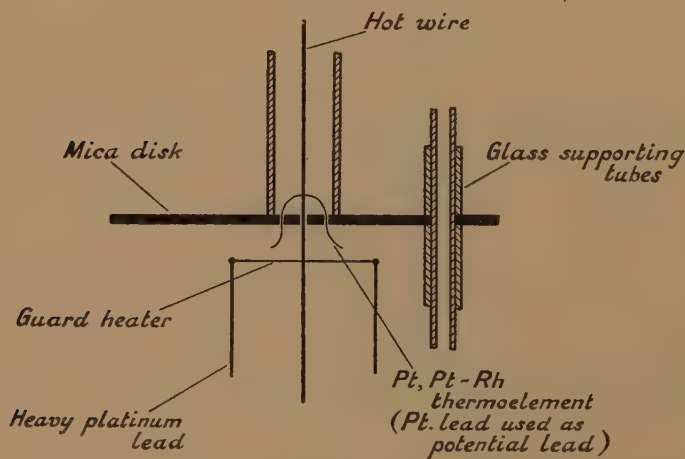
Two thermoelements, consisting of platinum and platinum-rhodium wire, are welded to the hot wire, one close to each end of the experimental region. The platinum leads of the thermoelements serve also as potential leads. Two auxiliary heaters of platinum wire are welded to the hot

wire close to the thermoelements, but just outside the experimental region.

The temperature of the tube surrounding the hot wire is measured by means of thermoelements attached to its outer surface, whilst the mean temperature of the hot wire itself is obtained from its resistance and the temperature at the ends by means of the two thermoelements.

The apparatus is enclosed in a glass envelope which is wound with copper foil, asbestos paper and nichrome ribbon. To reduce heat loss the apparatus is lagged with asbestos tape. The current and thermoelement leads are brought out of the apparatus *via* a wax seal. The lower end of the envelope is water-cooled to prevent heat conduction down to the seal.

Fig. 2.



A sectional view of the apparatus is given in fig. 1, while the detail at the lower end of the hot wire is presented on a larger scale in fig. 2.

The important dimensions of the apparatus are given below :—

Dimensions of Apparatus.

Diameter of hot wire	= 0.0500 cm.
Diameter of wire used as potential and the thermoelement leads	= 0.02 cm.
Distance between potential points	= 29.76 cm.
Resistance of hot wire at 0° C.	= 0.14659 ohm.
Internal diameter of tube surrounding hot wire	= 0.446 cm.
External diameter	= 0.652 cm.

The dimensions were measured at room-temperature except where otherwise indicated.

Experimental Procedure.

In carrying out an experiment the procedure is to exhaust the apparatus and fill with a sample of the gas under reduced pressure. Experiments carried out over a range of pressures from atmospheric down to about 5 cm. of mercury showed that the values of the conductivity obtained were independent of the pressure. This independence of conductivity of pressure can be regarded as strong evidence of the absence of convection.

The elimination of end loss of heat from the hot wire is a more difficult matter. It involves the adjustment of the currents flowing through the auxiliary heaters attached to each end of the hot wire just outside the experimental region, until the temperature at these points is as nearly as possible brought into equality with the mean temperature of the wire. The latter is determined from a knowledge of the current and potential drop in the wire.

In practice, it is found difficult to effect exact compensation for end conduction. So a separate series of experiments was carried out in which the ends of the hot wire were kept at temperatures differing by a few degrees from the mean; the object being to obtain an estimate of the effect on the apparent conductivity when the mean temperature of the wire differed by a definite amount from the temperatures at the potential points. It was found that, with air, the difference caused in the apparent conductivity by raising or lowering the temperature of the ends 1°C . above or below the mean temperature of the wire did not exceed 0.5 per cent. As the difference between end and mean temperatures was not more than 0.8°C . in any of the experiments, the results of which are given in this paper, any error from this source may be neglected.

In connexion with the readings of the thermoelements attached to the hot wire, it was found that the e.m.f. measured showed a variation when the current through the hot wire was reversed. This is due to the fact that it is impossible to locate both the platinum and the platinum rhodium limbs of the thermoelement at precisely the same spot on the heated wire. Since they are displaced by a minute distance, a potential difference is set up between the two points of attachment by the passage of the current. The mean of the readings before and after reversal of the current, however, gives the true temperature of the hot wire at the point of attachment.

The thermal emissivity of the platinum wire was determined by mounting the wire in front of a linear thermopile and heating it electrically. One half of the wire was painted with a thin coating of a black paint of 95 per cent. emissivity. The temperature of the bright and blackened portions was obtained by measuring the potential drop between two points on each.

In the test the temperature of the blackened portion was made equal to that of the bright part by adjustment of current. The thermal emis-

sivity of the platinum was 0.11 of that of a perfect black body at the same temperature.

Correction for Radiation.

The only correction necessary in these experiments was for the heat loss from the hot wire by radiation. For the purpose of calculating the radiation transfer, the surface of the glass tube was assumed to be a perfect emitter. The radiation transfer amounted to approximately 6 per cent. of the total heat transfer at 300° C. in the case of air.

Typical Experiment.

We give below the results of a typical experiment. Thermocouples 1 and 2 are attached to the hot wire and couples 3-6 give the temperature of the outer surface of the glass tube surrounding the hot wire.

Results of Experiment on "Freon."

Pressure = 600 mm. Hg.

Thermocouple.	Reading (microvolts).		Mean.	Temperature °C.
	Before reversal.	After reversal.		
1	1818	1757	1787	240.8
2	1827	1744	1785	

Current through hot wire = 1.802 amp.

Potential drop along hot wire = 0.5081 volt.

Thermocouple.	Reading (microvolts).	Mean reading.	Mean temperature °C.
3	1363	1361	191.6
4	1372		
5	1362		
6	1348		

Resistance of hot wire = 0.28197 ohm.

$$T_{pt} = 1740.6 (0.28197 - 0.14659) \\ = 234.6,$$

i. e., since $\delta = 1.50$, mean temperature of hot wire = 240.6° C. Heat generated per cm. length

$$= \frac{0.5081 \times 1.802}{4.186 \times 29.82} = 0.00734 \text{ cal/sec.}$$

Temperature drop through glass = 0.2°C .

Temperature drop through gas = $240.6 - 191.8 = 48.8^{\circ}\text{C}$.

Radiation transfer = $0.00054\text{ cal/sec./cm. length}$.

Transfer by conduction = $0.00680\text{ cal/sec./cm. length}$.

$$K = \frac{0.00680 \times 2.186}{2\pi \times 48.8} = 0.0000485\text{ cal/sec./cm.}^{\circ}\text{C. at a mean temperature of } 216^{\circ}\text{C.}$$

The coefficients of expansion of platinum and of Pyrex glass have been taken as 8.9×10^{-6} and 3×10^{-6} respectively.

Results.

The results obtained are set out below.

TABLE I.
Thermal Conductivity, K , in $\text{cal/cm./sec./}^{\circ}\text{C}$.

Air.		Carbon dioxide.		Freon.	
Temperature $^{\circ}\text{C}$.	$K \times 10^7$.	Temperature $^{\circ}\text{C}$.	$K \times 10^7$.	Temperature $^{\circ}\text{C}$.	$K \times 10^7$.
60.3	691	66.2	495	33.3	245
60.6	695	68.6	476	82.5	305
64.4	691	69.0	480	82.6	306
80.8	722	133.6	621	132.2	376
81.0	720	134.7	619	133.0	370
81.4	717	172.6	687	216.2	485
106.7	765	207.5	761		
111.0	788	243.7	847		
111.5	776	290.3	939		
130.3	820	292.4	910		
131.1	809				
155.4	857				
155.8	845				
156.7	871				
169.8	893				
202.8	941				
203.5	943				
205.3	945				
238.9	1001				
245.1	1012				
257.0	1016				
299.5	1103				
309.8	1116				
309.9	1100				

The results were smoothed graphically, and Table II. contains the smoothed values.

TABLE II.
Smoothed Values of K.

Temperature °C.	Thermal conductivity. $K \times 10^7$.		
	Air.	Carbon dioxide.	Freon.
0	(576)	(352)	(198)
50	(668)	(451)	264
100	759	550	330
150	848	649	397
200	934	748	463
250	1016	847	—
300	1096	945	—

A linear variation of K with temperature is found to represent the experimental values quite well for carbon dioxide and freon. The values enclosed in parentheses are extrapolated.

Discussion of Results.

An interesting relationship between the thermal conductivity, K, the viscosity, η , and the specific heat at constant volume, C_v , of a gas is indicated by the kinetic theory of gases.

It takes the form

$$K/\eta C_v = \epsilon,$$

where ϵ is a numerical factor, as to the value of which the theory is unable to give precise information.

No advance has been made since 1913, when Eucken derived the formula

$$\epsilon = (9\gamma - 5)/4,$$

in which γ is the ratio of the specific heats of the gas.

The thermal conductivity data which have hitherto been available for checking this formula are restricted to ordinary room-temperatures. In this region there is satisfactory agreement between the experimental and Eucken's value of ϵ for the majority of gases. It is of interest to compare the values of ϵ at various temperatures, derived from the present experiments, with those indicated by the formula.

Modern determinations have given the viscosity of air and carbon dioxide over the required temperature range (0° to 300° C.) with considerable

accuracy, and the position with regard to the specific heat at constant volume is nowadays equally satisfactory as there is abundant evidence that specific heats derived from spectroscopic data are substantially correct. If, then, we take the values of the thermal conductivity given in the present paper, we find that, in the case of air, ϵ increases from 1.96 to 2.06, as the temperature varies from 0° to 300° C. For carbon dioxide, ϵ is almost independent of temperature, increasing from 1.68 to 1.70 over the same temperature range.

On the other hand, the corresponding values derived from Eucken's equation are 1.91 and 1.85 for air, and 1.69 and 1.48 for carbon dioxide.

Whilst it might not be desirable at this stage to lay undue stress on the numerical values of the discrepancies between the values of ϵ deduced from experimental observations and those calculated from Eucken's formula, it might be noted that the experimental data indicate a small increase in ϵ with temperature, whereas the formula gives a decrease. The discrepancy between theory and experiment would appear to be too large to be accounted for by possible errors of observation.

The value of ϵ is a point of considerable interest and its more accurate determination might be justified. For further work it would be advisable to modify some of the details of the apparatus. In particular, the substitution of a platinum tube for the glass tube surrounding the hot wire suggests itself.

Acknowledgments.

We desire to thank Mr. A. R. Challoner for help with the design and assembly of the apparatus, and with the taking of the observations.

VII. *On the Neutron-Proton Scattering.*

By Prof. K. C. KAR, D.Sc., and D. BASU, B.Sc.*

[Received June 16, 1938.]

Introductory.

BEFORE investigating the problem of neutron scattering by proton or proton scattering by neutron, we think it necessary to discuss the closely related problem of neutron-proton combination forming deuteron. In both the problems the question of interaction potential is of great importance.

A study of the variation of mass defect with the atomic number shows that the nuclear binding energies of lighter nuclei, such as heavy hydrogen, *i. e.*, deuteron and helium, *i. e.*, α -particle, are as high as 2.15 mv., *i. e.*, 1.07 mv. per particle and 27.7 mv., *i. e.*, 6.92 mv. per particle respectively, whereas for elements of higher atomic number the mass defect or the binding energy is not correspondingly high, but only increases linearly with the number of particles in the nucleus. Thus, for example, the binding energy per particle for nuclei with atomic number round 30 and having the highest packing fraction, is only 8.5 mv. This strongly suggests that the combinations 1 neutron-1 proton (*i. e.*, deuteron) and 2 neutrons-2 protons (*i. e.*, α -particle) are highly stable and that the heavier nuclei are made up of these compound particles. Thus a heavy nucleus may be considered as a small quantity of liquid or solid, consisting of these compound particles. The excess of energy 8.5-6.9, *i. e.*, 1.6 mv. for $z \sim 30$, is evidently due to attractive potential energy of van der Waals type, owing to the presence of the neighbouring particles.

Again, the binding energy of the α -particles is thirteen times greater than the binding energy of the deuteron although the number of particles is only double. Thus it follows that the former combination is fully saturated, whereas the latter is only partly so. Therefore we are justified in assuming after Majorana that in deuteron there is saturation with respect to position, but not with respect to the spin. Or, in other words, on coming close to a proton, a neutron loses an electron which goes over to the proton converting it to a neutron, and so the positions of the neutron and proton are interchanged, but it makes no difference whether the spins of neutron and proton are parallel or antiparallel. Wigner, however,

* Communicated by the Authors.

assumes the interaction potential to be of the ordinary type, but it does not explain saturation.

Binding Energy of Deuteron.

Referring the motion to the centre of mass of the interacting neutron and proton, and assuming the potential to be spherically symmetrical, we have for the wave-equation of the neutron initially at $\mathbf{r}(r, \theta, \phi)$ and after exchange at $-\mathbf{r}(r, \pi - \theta, \pi + \phi)$

$$\frac{h^2}{4\pi^2 M} \Delta \chi(\mathbf{r}) + E \chi(\mathbf{r}) = V(r) \chi(-\mathbf{r}), \text{ Majorana, } \quad . \quad . \quad . \quad (1)$$

$$\frac{4\pi^2 M}{h^2} \Delta \chi(\mathbf{r}) + E \chi(\mathbf{r}) = V(r) \chi(\mathbf{r}), \text{ Wigner, } \quad . \quad . \quad . \quad (2)$$

where M is the mass of the proton or neutron being equal to double the effective mass, and E the binding energy being negative. Now, as

$$\begin{aligned} \chi(-\mathbf{r}) &= R(r) P_l^m(\cos(\pi - \theta)) e^{im(\pi + \phi)} \\ &= (-1)^l R(r) P_l^m(\cos \theta) e^{im\phi}, \end{aligned}$$

and

$$\chi(\mathbf{r}) = R(r) P_l^m(\cos \theta) e^{im\phi}$$

We have from (1)

$$\frac{h^2}{4\pi^2 M} \left\{ \frac{d^2}{dr^2} (rR) - \frac{l(l+1)}{r^2} (rR) \right\} + E \cdot (rR) = (-1)^l V(r) \cdot (rR). \quad . \quad (3)$$

It is evident that Wigner equation (2) leads to the above equation except the factor $(-1)^l$ on the right-hand side. For deuteron at the ground state $l=0$; so (3) reduces to

$$\frac{h^2}{4\pi^2 M} \cdot \frac{d^2}{dr^2} (rR) + E \cdot (rR) = V(r) \cdot (rR) \quad . \quad . \quad . \quad (3.1)$$

Thus for the ground state there is no difference between Majorana and Wigner hypotheses. Now in order to proceed further with the wave-equation (3.1) we must assume a definite form for the potential $V(r)$. Four forms of potential have been suggested so far, viz., (a) rectangular, in which $V(r) = -V_0(\text{const.})$ for $r < a$ and $=0$ for $r > a$; (b) $V = -V_0 e^{-\alpha r}$;

(c) $V = -V_0 e^{-\alpha r^2}$; and (d) $V = -\frac{4V_0}{(1+e^{-\alpha r})(1+e^{\alpha r})}$, taken by Wigner*.

Of the four forms the first is evidently only a simplified picture of the potential field. With the second and fourth forms of the potential rigorous solutions of the differential equation (3.1) are known. It is,

* Wigner, Phys. Rev. xliii. p. 252 (1933).

On substituting the above value of α in (5.1) we get

$$V_0 = 5.023 \times 10^{-5} \text{ erg} = 31.57 \text{ mv.} \quad (5.4)$$

In this way for any given value of s we can find the corresponding values of the parameters α and V_0 of the interaction potential. In order to determine s , α , and V_0 uniquely we require yet another equation connecting them. This we have from the energy equation. Thus from (3.1) we have for the binding energy

$$E = -\frac{\hbar^2}{4\pi^2 M} \int_0^\infty \frac{d^2}{dr^2} (rR) \cdot rR \cdot dr - V_0 \int_0^\infty e^{-\alpha r} R^2 r^2 dr, \quad (6)$$

and from (4)

$$= -\frac{\hbar^2 C^2}{4\pi^2 M} \int_0^\infty \frac{d^2 J_s}{dr^2} J_s dr - V_0 C^2 \int_0^\infty e^{-\alpha r} J_s^2 dr. \quad (6.1)$$

Now from (3.2) and (3.4) we have

$$\frac{dz}{dr} = -\frac{\alpha}{2} \cdot z,$$

and so

$$\frac{d^2 J_s}{dr^2} = \frac{\alpha^2}{4} z \left(\frac{d^2 J_s}{dz^2} z + \frac{dJ_s}{dz} \right),$$

and also

$$e^{-\alpha r} = \frac{\alpha^2 \hbar^2}{16\pi^2 M V_0} \cdot z^2.$$

Therefore from (6.1), denoting differentiation with respect to the argument by dash

$$E = -\frac{\hbar^2 C^2}{4\pi^2 M} \cdot \frac{\alpha}{2} \int_0^{z_0} J_s (z J_s'' + J_s') dz - \frac{\hbar^2 C^2}{4\pi^2 M} \cdot \frac{\alpha}{2} \int_0^{z_0} J_s^2 z dz, \quad (6.2)$$

where

$$z_0 = \frac{4\pi(MV_0)^{\frac{1}{2}}}{\alpha \hbar},$$

and in which the averaging factor is determined from the condition

$$C^2 \int_0^\infty J_s^2 dr \equiv \frac{2C^2}{\alpha} \int_0^{z_0} J_s^2(z) \frac{dz}{z} = 1. \quad (6.3)$$

Now from the well-known recurrence formulæ of Bessel functions

$$J_{s-1}(z) + J_{s+1}(z) = \frac{2s}{z} J_s(z),$$

$$J_s'(z) = \frac{s}{z} J_s(z) - J_{s+1}(z),$$

we get

$$\begin{aligned} J_s'(z) &= \frac{1}{2}\{J_{s-1}(z) - J_{s+1}(z)\} \\ J_s''(z) &= \frac{1}{2}\{J_{s-1}'(z) - J_{s+1}'(z)\} \\ &= \frac{1}{4}\{J_{s-2}(z) - 2J_s(z) + J_{s+2}(z)\}. \end{aligned}$$

With the help of the above (6.2) is transformed to

$$\begin{aligned} E = \frac{\alpha h^2 C^2}{8\pi^2 M} \left\{ \frac{1}{2} \int_0^{z_0} J_s^2(z) z dz + \int_0^{z_0} J_{s+1}(z) J_s(z) dz - s \int_0^{z_0} J_s^2(z) \frac{dz}{z} \right. \\ \left. - \frac{1}{4} \int_0^{z_0} J_{s+2}(z) J_s(z) z dz - \frac{1}{4} \int_0^{z_0} J_{s-2}(z) J_s(z) z dz \right\} - \frac{\alpha h^2 C^2}{8\pi^2 M} \int_0^{z_0} J_s^2(z) z dz, \end{aligned} \quad (6.4)$$

where the last term gives the average potential energy and the rest the average kinetic energy of the deuteron.

Integrating (6.3) graphically we get $C = 2.336 \times 10^6$ for $s=1$, $\alpha = 4.576 \times 10^{12}$ and $V_0 = 5.023 \times 10^{-5}$. With the above set of values we find from (6.4), after integrating by graphical method, $-E = 2.163$ mv., which is the experimental value of binding energy (*vide* (5.2))*. It may be noted that for $s=1.5$, we find from (6.4), on proceeding exactly as before, $-E = 2.216$, which is considerably greater than the experimental value. Thus it follows that for deuteron $s=1$, and α and v_0 are given in (5.3) and (5.4).

Scattering of Neutron by Proton.

The wave-equations of the neutron within and without the potential field, the motion being referred to the centre of mass, are respectively

$$\left. \begin{aligned} \Delta \chi + k^2 \left(1 - \frac{V}{E}\right) \chi &= 0, \\ \Delta \chi_0 + k^2 \chi_0 &= 0, \end{aligned} \right\} \quad \text{.} \quad (7)$$

where

$$k = \frac{\pi M v}{h} \quad \text{and} \quad E = \frac{1}{2} M v^2.$$

Proceeding as before †, we get for the first order scattering function

$$\lambda_1 \chi_1(r_2) = -\frac{1}{4\pi} \frac{k^2}{E} \int V(r_1) \chi_0(r_1) \frac{e^{ikr_{12}}}{r_{12}} d\tau_1, \quad \text{.} \quad (8)$$

where

$$\chi_0 = \frac{1}{\sqrt{v}} e^{ikx_1}, \quad \text{.} \quad (8.1)$$

* The latest value obtained by Chadwick (Proc. Roy. Soc. cxliii. p. 366 (1937)) is $2.170 \pm .04$ mv.; see also Bethe, Phys. Rev. liii. p. 313 (1938).

† K. C. Kar, M. Ghosh, and K. K. Mukherjee, Phil. Mag. xxiv. p. 968 (1937).

since the incident wave is supposed to be moving in the direction of the x -axis. On integrating (8) as before *, we get

$$\lambda_1 \chi_1(r_2) = - \frac{2k \operatorname{cosec} \frac{\theta}{2}}{Mv^2} \frac{e^{ikr_2}}{\sqrt{ur_2}} \int_{r_0}^{\infty} \sin k'r V(r) r dr, \quad \dots \quad (8.2)$$

where

$$k' = 2k \sin \frac{\theta}{2}$$

Or, we have

$$\lambda_1 \chi_1(r) = \frac{\operatorname{cosec}^2 \frac{\theta}{2}}{Mv^2} \cdot \frac{e^{ikr}}{\sqrt{v \cdot r}} F(r_0), \quad \dots \quad (9)$$

where

$$F(r_0) = -k' \int_{r_0}^{\infty} \sin k'r V(r) r dr, \quad \dots \quad (9.1)$$

and using the experimental law of potential

$$= k' V_0 \int_{r_0}^{\infty} e^{-\alpha r} \sin k'r r dr, \quad \dots \quad (9.2)$$

and integrating

$$= k' V_0 e^{-\alpha r_0} \left\{ \frac{\alpha \sin k'r_0 + k' \cos k'r_0}{\alpha^2 + k'^2} r_0 + \frac{(\alpha^2 - k'^2) \sin k'r_0 + 2\alpha k' \cos k'r_0}{(\alpha^2 + k'^2)^2} \right\}, \quad \dots \quad (9.3)$$

where r_0 is the critical approach of the incident neutron.

The relative intensity of scattering is

$$I_s = \frac{\operatorname{cosec}^4 \frac{\theta}{2}}{(Mv^2)^2} \cdot \{F(r_0)\}^2. \quad \dots \quad (10)$$

Critical Approach.

Using the boundary conditions

$$\left. \begin{aligned} \frac{d}{dr}(\chi_0 + \lambda_1 \chi_1) &= 0, \\ \chi_0 + \lambda_1 \chi_1 &= 0, \end{aligned} \right|_{r=r_0} \quad \dots \quad (11)$$

at $r=r_0$, and proceeding as in the previous paper (*l. c.*), we get for the critical approach

$$\tan kr_0 = kr_0 \quad \dots \quad (11.1)$$

* K. C. Kar, Phil. Mag. xxiv. p. 972 (1937); *vide* Eq. (4).

and

$$r_0 = \frac{V_0 \operatorname{cosec}^2 \frac{\theta}{2}}{Mv^2} \sin \left\{ 2\rho \left(\sin \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right) \right\} \cdot G(r_0),$$

where

$$\rho = 1.35$$

and

$$G(r_0) = k' \int_0^\infty e^{-\alpha r_0} \sin k'(r-r_0) r dr, \quad \dots \dots (11.3)$$

and integrating

$$= k'^2 e^{-\alpha r_0} \left\{ \frac{2\alpha}{(k'^2 + \alpha^2)^2} + \frac{r_0}{k'^2 + \alpha^2} \right\} \dots \dots (11.4)$$

On substituting the above value of $G(r_0)$ in (11.2) we get

$$r_0 = \rho \cdot \frac{8\pi^2 M V_0}{h^2} e^{-\alpha r_0} \left\{ \frac{2\alpha}{(\alpha^2 + k'^2)^2} + \frac{r_0}{\alpha^2 + k'^2} \right\} \left(\sin \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right). \quad (11.5)$$

First Approximation.

k is of the same order as, and is slightly less than, α ; so that $k' < \alpha$ at ordinary angles of scattering and therefore $k'^2 \ll \alpha^2$. Hence, neglecting k'^2 , we get from (11.5)

$$r_0 = \rho \cdot \frac{8\pi^2 M V_0}{h^2 \alpha^3} e^{-\alpha r_0} (2 + \alpha r_0) \left(\sin \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right). \quad \dots \dots (12)$$

It is evident from the above that $r_0 \rightarrow 0$ as $\theta \rightarrow 0$ or π . At other angles it has a small value. For $\theta = 90^\circ$, (12) gives $r_0 = 3.380 \times 10^{-13}$ cm. For such a small value of r_0

$$\sin k' r_0 \rightarrow 0, \quad \cos k' r_0 \rightarrow 1.$$

Hence from (9.3)

$$F(r_0) = \frac{k'^2 V_0 e^{-\alpha r_0}}{\alpha^3} (2 + \alpha r_0). \quad \dots \dots (13)$$

Therefore from (10) the relative intensity becomes

$$I_s = \frac{16\pi^4 M^2 V_0^2}{h^4 \alpha^6} e^{-2\alpha r_0} (2 + \alpha r_0)^2, \quad \dots \dots (14)$$

being independent of the angle of scattering. Thus the scattering is spherically symmetrical when referred to the centre of mass. This is generally supported by the experiments of Chadwick* and others. It is interesting to note that the intensity is also independent of the

* Chadwick, Proc. Roy. Soc. A, cxlii. p. 1 (1933); Kurie, Phys. Rev. xlv. p. 463 (1933); Monod-Herzen, *J. de Phys. et Rad.* v. p. 95 (1934).

incident velocity of the neutron. This should be so, since the intensity or the relative probability of scattering is independent of the scattering angle. So, it seems, the two effects are interdependent.

Second Approximation.

On retaining the higher order terms we have, instead of (14),

$$I_s = \left(\frac{4\pi^2 M V_0}{h^2 \alpha^3} \right)^2 e^{-2\alpha r_0} \left[(2 + \alpha r_0) \cos k' r_0 + \frac{\alpha}{k'} (1 + \alpha r_0) \sin k' r_0 \right]^2. \quad (15)$$

As (15) depends on k' , it is evident that to a second approximation the intensity depends both on the angle of scattering and the incident velocity.

Now, as $\frac{\sin k' r_0}{k'}$ is independent of k' since $k' r_0 \ll 1$, the variation of intensity with the angle of scattering is determined by the factor $\cos k' r_0$ in (15).

On looking back to (12) we find that

$$k' r_0 \propto \sin \frac{\theta}{2} \left(\sin \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right),$$

so that $\cos k' r_0$ attains the maximum value 1 for $\theta = 0$ or π , and for other angles it is less than unity. Thus the intensity of scattering first decreases with the increase of angle, but at large angles it again increases with the increase of the angle of scattering. This appears to be in accord with the experiments of Harkins * and others, who find larger number of protons at small angles with respect to the incident neutron than at large, and consequently larger number of neutrons at large angles than at a small.

In conclusion, it may be remarked that usually the cross-section of scattering is determined by $4\pi I_s$, but as in the present method the critical approach is known independently (*vide* (12)), we may define the cross-section as $\pi(2r_0)^2$, where r_0 is the critical approach for 90° scattering. However, we find from (12) and (14) that the ratio of the two cross-sections is

$$\frac{4\pi r_0^2}{4\pi I_s} = 4\rho^2 \left(\sin \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right)^2 \sim 2684 \sim \frac{1}{4}.$$

Physical Laboratory,
Presidency College, Calcutta, India.
April 1938.

* Harkins, Gans, Kamen, and Newson, Phys. Rev. xlvii. p. 511 (1935).

VIII. The Elementary Particle.

By Professor W. WILSON, F.R.S., and Miss J. CATTERMOLÉ, Ph.D.*

[Received October 25, 1938.]

IN macrophysical theory (Einstein's special theory of relativity) the motion of a particle conforms to the equation

$$p_x^2 + p_y^2 + p_z^2 - m^2 c^2 + m_0^2 c^2 = 0, \quad . \quad . \quad . \quad (1)$$

in which p_x , p_y , and p_z are its components of momentum referred to rectangular axes of coordinates, m and m_0 are its mass and rest mass respectively, and c is the velocity of light in free space. This equation is, of course, valid in any field of force if we may be permitted to ignore the refinements of Einstein's general theory of relativity. The problem of a particle (*e. g.*, electron) of microphysical dimensions is approached by regarding each term in (1) as an operator directed to some function, ψ , which we may regard as, in some sense, describing the particle. Thus instead of (1) we have the statement

$$(p_x^2 + p_y^2 + p_z^2 - m^2 c^2 + m_0^2 c^2) \psi = 0. \quad . \quad . \quad . \quad (2)$$

The individual operators are

$$\left. \begin{aligned} p_x &= (\hbar/2\pi i) \partial / \partial x, \\ p_y &= (\hbar/2\pi i) \partial / \partial y, \\ p_z &= (\hbar/2\pi i) \partial / \partial z, \\ \text{and} \quad mc &= -(\hbar/2\pi i c) \partial / \partial t \end{aligned} \right\} . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Therefore

$$\nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \frac{4\pi^2 m_0^2 c^2}{\hbar^2} \psi = 0. \quad . \quad . \quad . \quad . \quad (4)$$

When we give our attention to proper (*eigen*) solutions in which the time appears in the factor $e^{2\pi i \nu t}$, we may replace ψ in the last term by $-(1/4\pi^2 \nu^2) \partial^2 \psi / \partial t^2$ and thus obtain, since we must identify mc^2 with $h\nu$,

$$\nabla^2 \psi - \frac{1}{c^2} \left(1 - \frac{m_0^2}{m^2} \right) \frac{\partial^2 \psi}{\partial t^2} = 0,$$

or, since $m = m_0(1 - v^2/c^2)^{-\frac{1}{2}}$,

$$\nabla^2 \psi - \frac{1}{u^2} \frac{\partial^2 \psi}{\partial t^2} = 0, \quad . \quad . \quad . \quad . \quad . \quad (5)$$

* Communicated by the Authors.

in which the phase velocity, u , satisfies de Broglie's equation

$$uv=c^2. \quad (6)$$

Equations (2), (3), and (4) represent very considerable departures from classical mechanical theory, but they are not arbitrary ones: they are the inevitable outcome of expanding the parallelism between classical dynamics and geometrical optics *, and equation (5) does, in fact, account correctly for a great deal in the behaviour of electrons, as experiments on electron diffraction amply demonstrate. There are, however, phenomena associated with electrons of which this equation gives no hint, namely, those which are ascribed to an electronic magnetic moment and an associated angular momentum (spin) †. In order to find a theory which will endow the electron with a magnetic moment it appears to be necessary to regard the quadratic operator in equation (2) as emerging from more fundamental linear ones. Dirac ‡ represents it as a product of two linear factors

$$\{(\alpha_x p_x + \alpha_y p_y + \alpha_z p_z + \beta) + mc\} \times \{(\alpha_x p_x + \alpha_y p_y + \alpha_z p_z + \beta) - mc\} \psi = 0, \quad (7)$$

in which the α 's and β are matrices satisfying the relations

$$\alpha_x^2 = 1, \alpha_y^2 = 1, \alpha_z^2 = 1, \beta^2 = m_0^2 c^2, \alpha_x \alpha_y + \alpha_y \alpha_x = 0, \alpha_x \beta + \beta \alpha_x = 0, \text{ etc.} \S$$

According to Dirac's theory the electron has a spin momentum equal to $\pm \hbar/4\pi$. The factorization (7) does not, however, of itself yield a magnetic moment. To deduce this it is necessary to suppose the electron to be in an electromagnetic field and to express the p 's in terms of a more general momentum Π defined by

$$\Pi = p + eA/c, \quad (8)$$

where e is the charge on the particle and A is the vector potential (in suitable units) and to replace the product mc by $mc + eV/c$, where V is the scalar potential. These extended momenta Π are then identified with the operators (3). The new operators bring out an energy term $(eh/4\pi m_0 c) \times \text{curl } A$, so that a magnetic moment equal to $eh/4\pi m_0 c$ is assigned to the electron. It should be added that an electric moment also emerges; but this is ruled out as having no physical significance on the ground that it is purely imaginary. The success of Dirac's theory is

* W. Wilson, "The Origin and Nature of Wave Mechanics," 'Science Progress,' xxxii. no. 126, p. 209.

† S. Goudsmit and G. E. Uhlenbeck, *Naturwissenschaften*, xiii. p. 954 (1925); 'Nature,' cvii. p. 264 (1926).

‡ P. A. M. Dirac, "The Quantum Theory of the Electron," Proc. Roy. Soc. A, cxvii. p. 610 (1928) and cxviii. p. 351 (1928).

§ A. Eddington, "A Symmetrical Treatment of the Wave Equation," Proc. Roy. Soc. A, cxxi. pp. 524-542 (1928).

|| W. Wilson, Proc. Roy. Soc. A, cii. p. 481 (1922).

due to the association of a magnetic moment of a *whole* Bohr magneton with a spin momentum equal to *one-half* of the unit $\hbar/2\pi$.

The purpose of the present paper is to describe another way of deriving the quadratic operator (2) from linear ones. It is suggested by Kaluza's * unifying relativity theory. According to this theory the path of any particle, whether charged and in an electromagnetic field or not, is a geodesic in a 5-dimensional continuum, and we adopt the hypothesis that electrons, photons, and positrons are different aspects of one and the same thing and that their geodesics are *null* geodesics in Kaluza's continuum †. The photon is distinguished from the electron and positron, in our theory, only by the vanishing of the component of its momentum in the fifth dimension of the continuum. Now equation (4) becomes for a photon

$$\nabla^2\psi - \frac{1}{c^2} \frac{\partial^2\psi}{\partial t^2} = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

since m_0 vanishes, and we are led to suggest a more general equation to include the electron, positron and photon, namely,

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{\partial^2\psi}{\partial o^2} - \frac{1}{c^2} \frac{\partial^2\psi}{\partial t^2} = 0, \quad . \quad . \quad . \quad . \quad (10)$$

or, if we complete the symmetry by writing $w=ict$,

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{\partial^2\psi}{\partial w^2} + \frac{\partial^2\psi}{\partial o^2} = 0. \quad . \quad . \quad . \quad . \quad (10a)$$

Particular solutions of this equation will represent an electron, positron, or photon, as the case may be. When a definite momentum can be assigned to the particle the relationships between the phase and group (or particle) velocities are illustrated by the accompanying figure. The line OA represents the velocity, v , of the particle in space and OB its velocity, c , in the 4-dimensional continuum of space plus the dimension, o . When suitable observations have been made on the electron (photon or positron), the solution of (10), which describes it, will represent a plane wave having the phase velocity, c (in Kaluza's continuum), and some wave-length, λ_0 . The wave-length, λ , in space is, of course, the projection of this on OA as shown in the figure and

$$\lambda_0/\lambda = c/v = \cos \phi.$$

We have also

$$v/c = \cos \phi,$$

* Kaluza, "Zum Unitätsproblem der Physik," *Preuss. Akad. d. Wiss. Berlin*, liv. p. 966 (1921).

† J. W. Fisher has suggested that *all* particles travel in null geodesics in this continuum. *Proc. Roy. Soc. A*, cxxiii. p. 489 (1929).

so that the de Broglie relationship

$$uv=c^2$$

is satisfied. It may be noted, too, that

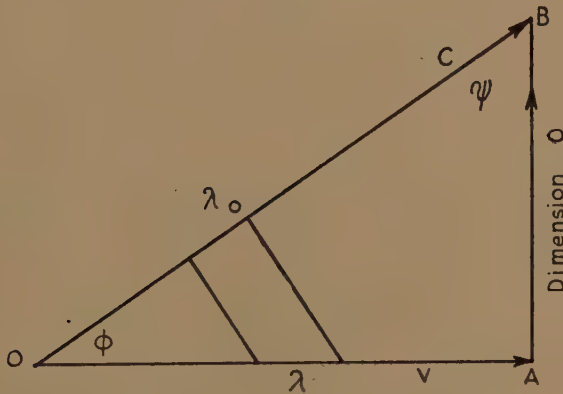
$$\cos \psi = \sin \phi = 1/\gamma,$$

where γ is the familiar coefficient of relativity theory, namely,

$$\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}.$$

Thus the velocity component parallel to the o axis, *i. e.*, parallel to AB, is $\pm c/\gamma$, and therefore the corresponding momentum component is $\pm mc/\gamma$ or $\pm m_0 c$, one sign being associated with the positron and the other with the electron. This momentum component vanishes for the photon since its rest mass is zero.

Fig. 1.



Now one of the fundamental things in the relativity theory we have adopted is the proportionality of the charge on the particle and its component of momentum in the extra dimension, o , of the continuum. Thus

$$K_0 e = m_0 c, \quad \dots \dots \dots (11)$$

K_0 being a constant.

This is simply the assumption * that conservation of charge is an aspect of conservation of momentum. Obviously when we regard $m_0 c$ as an operator and identify it, as we must do, with $(h/2\pi i) \partial/\partial o$ and substitute in equation (10) we get equation (4). It is clear, then, that equation (10) or (10 a) can represent either a positron or an electron or, in the special case where $m_0 = 0$, a photon.

* Flint and Wilson, Proc. Phys. Soc. 1. p. 340 (1938).

Now we have long been familiar with first-order differential equations which describe a photon. They are, in fact, Maxwell's equations for free space, the electric density and the stream vector being equated to zero. Four of them, which we may term the contravariant equations, when written in their most general form, are

$$\frac{1}{g^{\frac{1}{2}}}\frac{\partial(g^{\frac{1}{2}}F^{ab})}{\partial x_b}=0, \quad a, b=1, 2, 3, 4, \quad ; \quad . \quad . \quad . \quad . \quad (12)$$

in which the usual summation convention is implied. There is, of course, one of these equations for each value of a and therefore four in all. We shall take space-time to be Galilean, and therefore, with suitable co-ordinates, (12) becomes

$$\frac{\partial F^{ab}}{\partial x_b}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (12 a)$$

Associated with these contravariant equations are four covariant ones. We may represent them by the typical equation

$$\frac{\partial F_{ab}}{\partial x_c}+\frac{\partial F_{bc}}{\partial x_a}+\frac{\partial F_{ca}}{\partial x_b}=0, \quad . \quad . \quad . \quad . \quad . \quad (12 b)$$

(a, b, c) being any one of the four combinations of 1, 2, 3, 4 taken three at a time.

The corresponding covariant and contravariant components of F may be identified in a Galilean continuum when we employ rectangular coordinates. Therefore

$$F^{ab}=F_{ab}=-F_{ba}$$

and, of course, $F^{aa}=F_{aa}=0.$

With suitable units

$$F_{12}=H_z, \quad F_{23}=H_x, \quad F_{31}=H_y, \quad F_{41}=iE_x, \quad F_{42}=iE_y, \quad F_{43}=iE_z,$$

H and E being the magnetic and electric field intensities respectively. The presence of $i(=\sqrt{-1})$ is explained by the fact that we are using instead of the time, t , the variable w or $x_4(=ict).$

The waves described by the wave-equations representing a particle, *e. g.*, by Maxwell's electromagnetic equations in the case of a photon, are not physical entities in the sense that they are observable things. Their function is to enable us to calculate probabilities. In the case of a photon of known frequency, for example, the expression

$$\frac{1}{2\hbar\nu}(E^2+H^2)dx\,dy\,dz$$

is the probability that the photon is in the volume element $dx\,dy\,dz.$ E and H only approximate to their old-fashioned classical significance when we are dealing with enormous numbers of photons, as, for example, in the

long waves used in broadcasting. When we remember these things and also that the electric and magnetic field vectors have a significance, not only for photons, but even more obviously for electrons and positrons, we are naturally led to represent these latter by first-order differential equations, which are, in fact, Maxwell's equations as they would be in a continuum with an extra dimension. The electron, photon, or positron is represented, as we have seen, by a second-order equation (10 *a*), and the first-order Maxwell equations from which we can derive it are, in the first place, *five* contravariant ones, namely

$$\frac{\partial F_{\alpha\beta}}{\partial x_\beta} = 0, \quad \alpha, \beta = 1, 2, 3, 4, 5. \quad (13)$$

For example, the contravariant equation with $\alpha=1$ is

$$\frac{\partial F_{11}}{\partial x_1} + \frac{\partial F_{12}}{\partial x_2} + \frac{\partial F_{13}}{\partial x_3} + \frac{\partial F_{14}}{\partial x_4} + \frac{\partial F_{15}}{\partial x_5} = 0,$$

which may otherwise be written

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \frac{1}{c} \frac{\partial E_x}{\partial t} + \frac{\partial F_{15}}{\partial x_5} = 0,$$

so that we naturally make the identifications

$$\frac{\partial F_{51}}{\partial x_5} = \rho v_x'/c, \quad \frac{\partial F_{52}}{\partial x_5} = \rho v_y'/c, \quad \frac{\partial F_{53}}{\partial x_5} = \rho v_z'/c, \quad \frac{\partial F_{54}}{\partial x_5} = i\rho, \quad (13a)$$

v' being the velocity of convection of the electric density ρ .

There are *ten* covariant equations, *i. e.*, as many as there are combinations of five things taken three at a time. A typical one is

$$\frac{\partial F_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial F_{\beta\gamma}}{\partial x_\alpha} + \frac{\partial F_{\gamma\alpha}}{\partial x_\beta} = 0, \quad (13b)$$

in which α, β , and γ are any three numbers chosen from 1, 2, 3, 4, and 5. We regard the electron (or positron) rather in the way in which Schroedinger regarded it in his picture of the hydrogen atom *, as if it were spread over space, so that the integral $\iiint \rho dx dy dz$ over the spatial region occupied by the wave at any instant is equal to e the electronic charge; but instead of interpreting ρ as an actual charge density, we attach such a meaning to it that

$$\rho dx dy dz/e$$

is the probability that the electron is in the volume element $dx dy dz$. In this connexion we may remark that the fifth of the equations (13) is, in fact, equivalent to

$$\text{div}(\rho v') + \partial \rho / \partial t = 0. \quad (14)$$

* E. Schroedinger, 'Wave Mechanics,' Blackie and Son, 1928, p. 18.

To see this we differentiate the equation

$$\frac{\partial F_{i\beta}}{\partial x_\beta} = 0$$

by x_5 and replace each $\partial F_{i\beta}/\partial x_5$ by the appropriate expressions in (13 a).

From equations (13) we get

$$\frac{\partial^2 F_{\alpha\beta}}{\partial x_\gamma \partial x_\beta} = 0,$$

or

$$\frac{\partial}{\partial x_\beta} \left(\frac{\partial F_{\alpha\beta}}{\partial x_\gamma} \right) = 0,$$

and therefore by (13 b)

$$\frac{\partial}{\partial x_\beta} \left(\frac{\partial F_{\gamma\beta}}{\partial x_\alpha} + \frac{\partial F_{\alpha\gamma}}{\partial x_\beta} \right) = 0.$$

Consequently

$$\frac{\partial}{\partial x_\alpha} \left(\frac{\partial F_{\gamma\beta}}{\partial x_\beta} \right) + \frac{\partial^2 F_{\alpha\gamma}}{\partial x_\beta^2} = 0.$$

The first term vanishes by (13), and therefore

$$\frac{\partial^2 F_{\alpha\gamma}}{\partial x_\beta^2} = 0. \quad \dots \dots \dots (15)$$

It will be remembered that in these equations a summation with respect to β is implied and (15) has, consequently, exactly the same form as (10) or (10 a).

The function ψ of these equations may be any one of the components of the tensor F .

When we substitute F_{54} for $F_{\alpha\gamma}$ in (15) and differentiate with respect to x_5 we get

$$\frac{\partial^2 \rho}{\partial x_\beta^2} = 0. \quad (\text{Summation with respect to } \beta.)$$

If we confine our attention to proper solutions in which the time appears in the factor $e^{2\pi i v t}$ and remember that

$$m_0 c = \left(\frac{h}{2\pi i} \right) \frac{\partial}{\partial x_5},$$

we get

$$\nabla^2 \rho - \frac{1}{u^2} \frac{\partial^2 \rho}{\partial t^2} = 0,$$

so that ρ is propagated with the phase velocity, u , where $u = c^2/v$.

Of the covariant equations (13 b) one is the familiar equation usually written $\text{div } B = 0$, and three of them are the Maxwell equations containing the components of curl E . They are, in fact, identical with the four covariant equations which appear in the description of the photon.

Of the remaining six covariant equations three endow the electron with a magnetic moment. For example,

$$\frac{\partial F_{12}}{\partial x_5} + \frac{\partial F_{25}}{\partial x_1} + \frac{\partial F_{51}}{\partial x_2} = 0,$$

when differentiated by x_5 , becomes

$$\frac{\partial}{\partial x_1} \left(\frac{\partial F_{25}}{\partial x_5} \right) + \frac{\partial}{\partial x_2} \left(\frac{\partial F_{51}}{\partial x_5} \right) = \frac{\partial^2 F_{21}}{\partial x_5^2}.$$

When we replace $F_{12} = -F_{21}$ by H_z and make use of (13 a), we get

$$\frac{\partial}{\partial x} \left(\frac{\rho v_y'}{c} \right) - \frac{\partial}{\partial y} \left(\frac{\rho v_x'}{c} \right) = \frac{\partial^2 H_z}{\partial o^2},$$

or, finally,

$$\text{curl } (\rho v') = c \frac{\partial^2 H}{\partial o^2}. \quad (16)$$

Now, as we have seen, the operation $\partial/\partial x_5$ or $\partial/\partial o$ is equivalent to multiplication by $(2\pi i/\hbar)m_0 c$. Therefore

$$\text{curl } \rho v' = - \frac{4\pi^2 m_0^2 c^3}{\hbar^2} H.$$

If ω is the associated angular velocity

$$2\rho\omega = - \frac{4\pi^2 m_0^2 c^3}{\hbar^2} H. \quad (16 a)$$

The rotation ω necessitates a magnetic moment and a spin at the same time. If dM be the contribution to the magnetic moment due to the rotation in a small region, and dI the associated contribution to the moment of inertia,

$$dM = \frac{\rho\omega}{2\rho'c} dI, \quad (17)$$

where ρ' is the (rest) mass density corresponding to the electric density ρ . The ratio ρ/ρ' of charge to mass must be equal to e/m_0 by (11). Therefore

$$M = \frac{e}{2m_0 c} J, \quad (18)$$

where J is the angular momentum of the electron (positron).

It thus appears that the ratio M/J is the classical one. This is not so serious as at first sight it appears to be. The principle of conservation of momentum, which in our theory comprehends that of mass, energy and charge, requires that the creation of an electron (which happens, for example, when a β -particle is emitted in the radioactive disintegration of an atom) must be associated with that of a positron. We have, of

course, to assume that the positron does not leave the nucleus, or at all events does not leave its immediate neighbourhood. Such a hypothesis has indeed already been made*. This positron having an equal and opposite spin momentum to that of the electron has an equal magnetic moment directed in the same sense. This is all that is required (if we assign to the electron, as indeed we must do, the spin momentum $\hbar/4\pi$) to explain the so-called anomalous Zeeman effect, since the electron and positron together produce the full Bohr unit of magnetic moment.

All difficulties associated with the conservation of spin momentum are obviated by the positron, and probably also those associated with the conservation of energy, without the need of the neutrino. It should, however, be added that, while the theory sketched in this communication is not in conflict with the known facts about the gyromagnetic effect and the gyromagnetic anomaly, it does not (in so far as we have been able to develop it) account for them.

The three remaining covariant equations are

$$\frac{\partial F_{45}}{\partial x_\alpha} + \frac{\partial F_{5\alpha}}{\partial x_4} + \frac{\partial F_{\alpha 4}}{\partial x_5} = 0, \quad \alpha = 1, 2, 3.$$

It will suffice to consider the case $\alpha = 1$, and we shall write x for x_1 and replace x_4 by ict . We easily find on differentiating by x_5 or o , and making use of (13 a),

$$\frac{\partial \rho}{\partial x} + \frac{\partial(\rho v'_x/c)}{c \partial t} = - \frac{\partial^2 E_x}{\partial o^2}.$$

On multiplying by $K_0 c$ (cf. 11) we get

$$\frac{\partial(\rho' c^2)}{\partial x} + \frac{\partial}{\partial t}(\rho' v'_x) = -K_0 c \frac{\partial^2 E_x}{\partial o^2}$$

in which ρ' is the (rest) mass per unit volume. Euler's hydrodynamical equations suggest that we should replace the left-hand member of this equation by

$$2\rho'(v'_y \omega_z - v'_z \omega_y),$$

ω being the angular velocity in (16 a). Thus we obtain

$$2\rho'(v'_y \omega_z - v'_z \omega_y) = -K_0 c \frac{\partial^2 E_x}{\partial o^2},$$

and on substituting for ω from (16 a) and remembering that

$$\frac{\partial^2 E_x}{\partial o^2} = - \frac{4\pi^2 m_0^2 x^2}{\hbar^2} E_x,$$

$$\frac{2\rho'}{2\rho} \{-v'_y H_z + v'_z H_y\} = K_0 E_x,$$

* G. Beck, International Congress on Physics, London, 1934, p. 31; published by the Physical Society.

and since $K_0 = \rho'c/\rho$, we get, finally,

$$E_x + \frac{1}{c}(v_y'H_z - v_z'H_y) = 0,$$

or

$$f = E + \frac{1}{c}[v'H] = 0.$$

This result is to be expected since we are dealing with a particle which is moving in its own field and not subject to an external force. Hence

$$\iiint f \rho \, dx \, dy \, dz$$

must vanish.

IX. *A Note on the Method of Parallax and the Resolving Power of the Eye.*

By SUKHDEO BIHARI MATHUR, M.Sc., Hindu College,
University of Delhi, Delhi *.

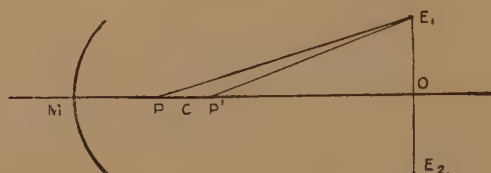
[Received June 16, 1938.]

SUMMARY.

The paper describes some observations made to determine the accuracy obtainable in ordinary laboratory experiments on the method of parallax and its relation with the resolving power of the eye.

IN elementary optical experiments on the determination of constants of mirrors and lenses the method of parallax is often employed. The

Fig. 1.



method consists, as is well known, in locating the position of the image by the help of another object brought in coincidence with it, the coincidence being judged by the eye. The purpose of the present note is to determine experimentally the extent of accuracy which is possible in this kind of experiment, and to correlate this with the resolving power of the eye.

To avoid the effects of any undue complications the experimental arrangement was taken to be of the simplest kind. A concave mirror of radius of curvature of about 20 cm. was fitted on a precision optical bench.

In fig. 1 M represents the mirror, P the position of the pin, and P' the image of P as formed in the mirror. The eye is placed at O and moved to and fro at right angles to the axis from E_1 to E_2 , and the position of the pin is adjusted till the parallax between P and P' disappears. We shall

* Communicated by Dr. D. S. Kothari.

denote the distance PO by D and the amplitude ($OE_1=OE_2$) of the swing of the eye by d .

The observations were taken in the following order : d was kept equal to 4, 6, 10, 16, and 20 cm., and D equal to 10, 30, 50, 70, and 90 cm. for each of the above values of d . Keeping d fixed at 4 cm., D was changed to 10, 30, 50, 70, and 90 cm. successively, and for each of these values of D , a set of twenty readings* was taken for the position P when parallax disappeared between P and P'. These observations were then repeated for other values of d , namely 6, 10, 16, and 20 cm.

In order to bring out the relation between the resolving power of the eye and the method of parallax let us again consider fig. 1. If P and P' denote the position of the pin and its image, then, when the eye is at E_1 , the angle subtended by PP' at the eye is

$$\alpha = \frac{PP' \sin \theta}{PE_1} = \frac{\Delta x \cdot d}{D^2 \left(1 + \frac{d^2}{D^2}\right)} \dots \dots \dots (1)$$

P, P' will appear separate to the eye so long as this angle is greater than the resolving power of the eye. If P and P' are so close that the angle PE_1P' becomes equal to or smaller than the resolving power of the eye, then P, P' will not appear separate, and as this angle will be all the smaller as the eye moves from E_1 towards O, it is obvious that the parallax between P and P' will disappear when the eye is moved through the range E_1OE_2 .

Let K denote the angle corresponding to the resolving power of the eye, then the distance PP', such that it subtends an angle K at E_1 , will be given by calling it ϵ ,

$$\epsilon = K \cdot \frac{D^2}{d} \left(1 + \frac{d^2}{D^2}\right) \dots \dots \dots (2)$$

We have now only to find the relation between ϵ and our observations of the position of P. The eye when moved to and fro from E_1 to E_2 will be unable to appreciate any parallax between P and P' when the distance between P, P' will be less than or equal to ϵ .

We shall assume here that all positions of the pin P such that the value of the separation PP' lies in the range 0 to ϵ are equally probable.

Let x denote the distance of P, measured from some fixed reference point, when the parallax between P and P' disappears, and let x_0 be the mean value of x . In our particular experiment x_0 will represent the position of C, the centre of curvature of the mirror. Then the mean

* These observations were taken off and on during the course of the last year.

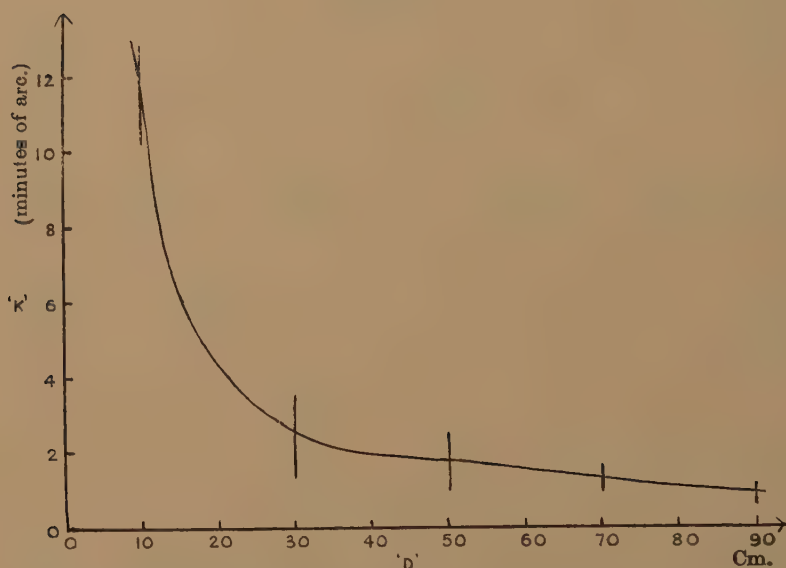
square deviation, i. e., the average of $(x-x_0)^2$, which we shall denote by $\overline{\Delta x^2}$, is

$$\overline{\Delta x^2} = \int_{x-x_0=-\frac{\epsilon}{2}}^{x-x_0=\frac{\epsilon}{2}} \frac{dx}{\epsilon} (x-x_0)^2 = \frac{\epsilon^2}{12}, \quad \dots \dots \dots (3)$$

or $\epsilon = \sqrt{12} \cdot (\overline{\Delta x^2})^{\frac{1}{2}}, \quad \dots \dots \dots (4)$

or $\sqrt{\overline{\Delta x^2}} = 0.29 \epsilon. \quad \dots \dots \dots (5)$

Fig. 2.



K represents the resolving power of the eye in minutes and
D represents the distance of the eye from the object.

The limits of integration are taken from $\frac{\epsilon}{2}$ to $-\frac{\epsilon}{2}$, for when the pin P is moved a little to one side of C the image P' will move by the same amount on the other side, and so when P moves by $\frac{\epsilon}{2}$ away from C, P' will be at a distance of $\frac{\epsilon}{2}$ on the other side of C, the separation PP' being equal to ϵ , the limit of resolution of the eye.

It must be mentioned that the factor connecting ϵ and $\overline{\Delta x^2}$ is to a certain extent arbitrary, as it depends upon the assumption we make regarding

the *a priori* probability of the distribution of observations in the interval $-\frac{\epsilon}{2}$ to $+\frac{\epsilon}{2}$. However, the factor will not be much different from $\sqrt{12}$, and we shall take it here to be so. Thus finally we have, using (2) and (4),

$$K = 1.2 \times 10^4 \frac{(\overline{\Delta x^2})^{\frac{1}{2}} \cdot d}{D^2 \left(1 + \frac{d^2}{D^2}\right)} \text{ minutes of arc.} \quad . \quad . \quad . \quad (6)$$

The results of our observations are shown in fig. 2, which represents K , the resolving power of the eye in minutes, plotted against D , the distance of the eye from the pin. The vertical lines indicate the spread in the value of K due to varying d . It will be seen that K attains the more or less constant value of about $40''$ when the distance D approaches the order of a metre. This is near enough the usual resolving power of the eye, which is taken to be about $1'$. For $D=10$ cm. K is abnormally large. This is easily explained, as for $D=10$ cm. the vision loses in acuity, because this distance is less than the least distance of distinct vision. Taken on the whole, our observations show that the accuracy obtainable in the method of parallax under ordinary laboratory conditions is limited by the resolving power of the eye itself.

As a check on the above statical observations were also taken, *i. e.*, the eye was kept fixed at the position E_1 , and the absence of parallax was judged with the eye fixed and without moving it from E_1 to E_2 . As is to be expected from the discussion given above it was found that the values of K obtained for different D and d from these statical observations were in fair agreement with those obtained when the eye was moved to and fro in the range E_1OE_2 .

It shows therefore that in the method of parallax the movement of the eye is not essential—it does not impart any increased accuracy. As is seen from equation (6) $\overline{\Delta x^2}$ depends inversely on d , and directly on D^2 , and thus for increasing the accuracy we should take d , *i. e.*, the perpendicular distance of the eye from the optical axis as large as is easily possible under the experimental conditions, and D should be kept small, nearly the least distance of distinct vision—about 25 cm.

I desire to express my thanks to Dr. D. S. Kothari for his interest in this work.

X. The Geometry of Discrete Vector Maps.

By D. M. WRINCH, M.A., D.Sc., Mathematical Institute, Oxford *.

[Received December 3, 1938.]

SUMMARY.

The geometry of finite and periodic point sets in S_1 and of their corresponding vector sets in S_2 is discussed. Methods are given for deriving from any S_2 set the complete class of S_1 sets corresponding. The application of these ideas to crystal analysis is indicated. It is shown that by means of this technique X-ray data may be used to discover crystal structures, not merely to test structures already devised.

1. Introduction.

It has long been recognized that a crystal may conveniently be represented by means of a discrete distribution of point intensities which represent positions of high or low density, and recent studies in crystal analysis ⁽¹⁻⁴⁾ have drawn attention to the possibility of a systematic geometrical formulation of this fact. Vector maps of such point intensity sets have been constructed and compared with vector diagrams ⁽⁵⁾ (*i. e.* projections and sections of vector maps) calculated from the amplitudes of structure factors obtained by X-radiation studies of actual crystals. Since these vector diagrams give a complete picture of all the experimental data it is necessary and sufficient for the testing of a proposed structure that there should be concordance between these diagrams and the vector maps of the proposed structure. But theoretical questions of great interest remain to be discussed ⁽⁶⁾. Given a vector map which is satisfied by a certain distribution in the crystal what, if any, alternative distributions also satisfy it? Since it would be of great importance to be able to use vector maps for the discovery of crystal structure the second question runs: Given a vector map, how can we find the complete set of distributions corresponding to it? It has so far not proved possible to answer either of these questions when the distribution is given in the form of a continuous electron density function ⁽⁷⁾. When the distribution is expressed in terms of a point intensity set, on the other hand, both questions can be answered in a simple manner. This communication indicates, in outline, the method by which this can be accomplished. For purposes of exposition

* Communicated by the Author.

considerations are confined to plane sets. The principles are, however, the same for sets in any number of dimensions. Since the projection of a vector map of a set of points is the vector map of the projection of the set, the interpretation of plane maps gives all that is required in practice.

In a few exceptional cases it may be found that, given one solution, it is possible to study how it can be modified without the vector map being changed. This fact is of importance in practice. In the event of a proposed structure passing the X-ray test (*i. e.*, yielding a vector map consistent with the experimentally obtained vector diagrams) we can then deduce a severely limited set of alternative structures, one of which must be correct even if it itself is not. However, the important point is that even when no structure at all has been proposed, the methods here introduced allow any vector point diagrams to be analysed into a finite number of alternative point sets (rarely more than a very few), one among which must represent the structure of the crystal to the same degree of accuracy as the vector diagrams represent the facts.

2. Vector Maps of Point Sets.

Let us take in space S_1 two points (1+2) and write 12 for the vector from 1 to 2. We then erect, at some point O in space S_2 , vectors corresponding to the vectors 11, 21, 12, 22 in S_1 , calling the free end points of these vectors in S_2 11, 21, 12, 22 respectively. Then these four points in S_2 constitute the vector map of the two points in S_1 . The points 11 and 22 lie at O. The points 21 and 12 are symmetrically placed with regard to O. Any two such points may be called a "pair," and each will be called the "partner" of the other. Lines joining pairs of points may be similarly called "line pairs," and so on.

If we now take a set of points in S_1 , (1+2+3... n), we can in the same way build the vector map

$$V(1+2+3 \dots n) = (11+22 \dots +nn) + (12+21) + (13+31) \\ + (n-1 \ n + n \ n-1) = \Sigma mm + \Sigma (mp)_2,$$

where $(mp)_2 \equiv mp + pm$. V thus comprises in all n^2 points in S_2 . n of these coalesce to make the origin a compound point of order n ; the rest form k point pairs $(mp)_2$, where $k = n(n-1)/2$, showing that every vector map has a centre of symmetry. In the case in which no vectors in S_1 are equal no point pairs in S_2 coalesce, and the number of such point pairs is then 3, 6, 10, 15, 21... according as n is 3, 4, 5, 6, 7... A number of point pairs not of this form indicates that there are no corresponding S_1 sets without at least one pair of equal vectors.

Vector maps can be built by a process of accretion. Thus

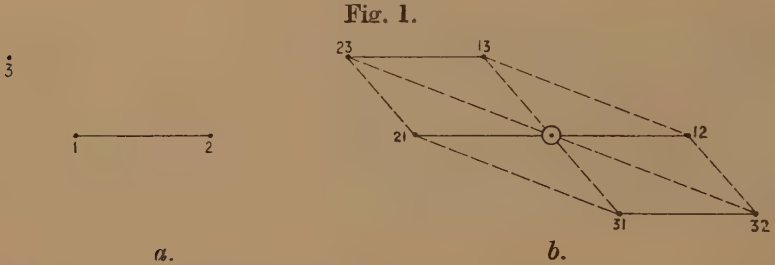
$$V(1+2 \dots +n+m) = V(1+2 \dots +n) + V(m) + m(1+2 \dots +n)_2.$$

We proceed with the vector map of 3 points (fig. 1 *a*), which may be written in the forms

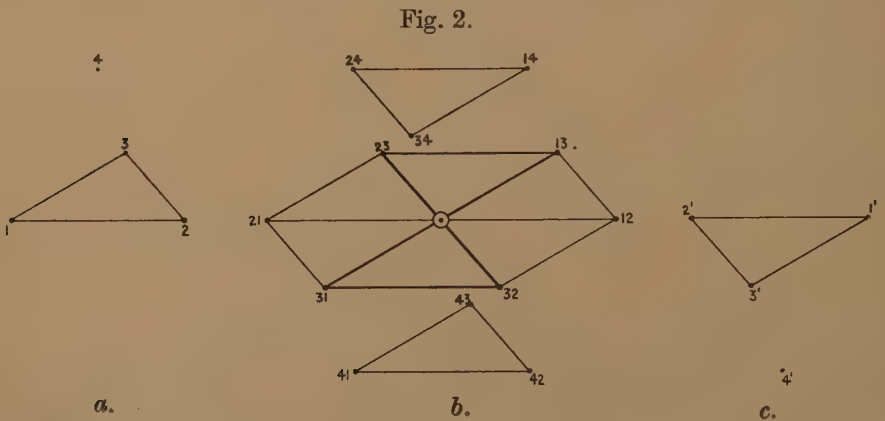
$$V(1+2+3)=(11+22+33)+(12)_2+(13)_2+(23)_2 \quad \dots \quad (1)$$

$$=(1+2)(1+2)+3(1+2)+(1+2)3+33. \quad \dots \quad (2)$$

The second formulation draws attention to the anatomy of this vector map, which is shown in fig. 1 *b*. For any set of points 1, 2, 3 the three



vectors 12, 23, 31 form a closed circuit, *i. e.*, $12+23+31=0$. This may also be written $31-32=21$. Thus, writing $31-32=3(1-2)=21$ we see (as shown in fig. 1 *b*) that the S_2 line from 32 to 31, namely $32-31$, is a parallel displacement of the S_1 vector 21. The points 32, 31 are in fact the image in 3 of the points 2, 1. Thus $V(1+2+3)$, as indicated in (2),



consists of $V(1+2)$ together with $V(3)$ (another point at the origin O), and in addition $3(1+2)$, the image of $1+2$ in 3 and $(1+2)3$ the images of 3 in 1 and in 2.

Treating next the vector map of four coplanar points $1+2+3+4$ (fig. 2 *a*) we notice that it comprises, in addition to the vector map of the three points $1+2+3$, a new point 44 at the origin, together with two

partner triangles, representing respectively the image in 4 of the triangle $1+2+3$ and the images of the point 4 in the three points of this triangle (fig. 2 b). If we replace each point 1, 2, 3, 4 by its partner with respect to any point in their plane we obtain a second set $1', 2', 3', 4'$ with the same vector map. Plainly, in this and in every case, whenever any S_1 has been shown to have a vector map V , then there is a second S_1 , the "partner" of the first, which also has this property. Further, when the point set is defined by coordinates in S_1 space, then V is invariant for any parallel displacement of these S_1 points. For any rotation V is rotated bodily by the same amount, and, in particular, rotation through two right angles (which results in the partner set) leaves V unchanged, since V has a centre of symmetry.

3. Vector Maps of Point Intensity Sets.

In this and all other cases under consideration in this paper we regard the S_1 and S_2 specifications as relating essentially to sets of points. There is no difficulty in extending this treatment to deal with the problem of the interrelations of S_1 and S_2 sets of point intensities. Thus if n be a point of intensity i_n , we specify an S_1 set as $\Sigma i_n n$, and defining the interactions of two point intensities $i_n n$ and $i_m m$ as $i_n i_m (nm)_2$, we have as the vector map of the S_1 point intensities

$$V(\Sigma i_n n) = \Sigma i_m^2 m m + \Sigma i_m i_p (mp)_2,$$

a vector map which belongs also to the distribution $\Sigma(-i_n n)$, in which the sign of every intensity has been changed throughout.

We notice the important facts that the intensity at O in S_2 is the sum of the squares of all the intensities in S_1 , and that the sum of all the intensities in S_2 is the square of the sum of all the intensities in S_1 . Evidently when we know the complete class of S_1 point sets corresponding to a single S_2 point set we have the ground plan of all the S_1 point intensity sets which can correspond to a given S_2 point intensity set *. Whether or not some or all or none of these can have intensities associated with the various points in such a way as to tally with the given intensities of the S_2 points is then an additional question. An n point intensity set in S_1 gives an n^2 point intensity set in S_2 . But the number of independent intensities in S_2 is never greater than $1 + \frac{1}{2}n(n-1)$, whereas the number of point intensities in S_1 is n . Thus unless there are adequate numbers of identical vectors in S_1 with different intensities the problem

* We assume at present that all points of the vector map are specified, including those points belonging to it which happen to have zero intensity. We do not here discuss the analysis of a vector point intensity map which, in addition to given entries with positive or negative intensities, may include unspecified points, provided these have zero intensities.

cannot be solved for every set of S_2 intensities. Thus even when it has been shown that an S_2 point set is the vector map of a set of points, an S_2 point intensity distribution with the same ground plan will not necessarily correspond to any point intensity set in S_1 . On the other hand, if there is an S_1 point intensity set corresponding to a given point intensity set in S_2 , this is unique apart from its partner set and the same sets with intensities of the opposite sign, if there are no identical vectors in the S_1 set. An example of a multiplicity of intensities associated with a given S_1 ground plan which has two or more identical vectors is discussed later.

4. *The Analysis of Vector Maps.*

We proceed to apply these principles to the problem of analysing an S_2 , *i. e.*, to the problem of finding which, if any, S_1 sets have the given

Fig. 3 (*a, b*).



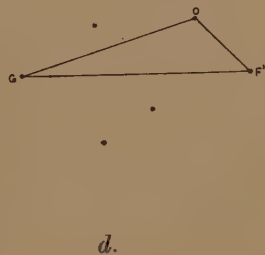
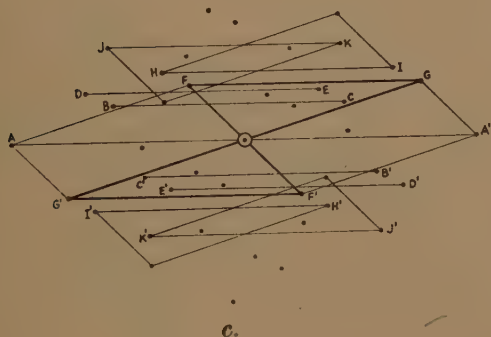
S_2 set as their vector map. It is convenient to talk of such an S_1 as a root of the S_2 and the points of such a set as radical points of S_2 .

Fig. 3 *b* represents an S_2 which was in fact constructed from the S_1 set pictured in fig. 3 *a*. We select any point pair in S_2 —say the furthest apart AA' . We then (in fig. 3 *c*) draw in all S_2 lines which are parallel displacements of AO (or of OA'), namely the 5 lines BC, DE, FG, GH, IJ , and the partner lines $B'C', D'E', F'G', H'I', J'K'$. We select arbitrarily one line and its partner, say $FG, F'G'$, and complete the hexagon $AFGA'F'G'$. The sides of the hexagon form three line pairs. This hexagon draws our attention to a pair of triangles FGO and $F'G'O$, each of which has as its vector map the six triangles which together make up the hexagon.

We now draw in an S_1 space the triangle $F'G'O$ as shown in fig. 3 *d*. We then examine each of the remaining lines to see whether a point exists which with the 2 points on the line gives an image either of the triangle $F'G'O$ or of the triangle FGO . We find in fact one point which with HI gives an image of $F'G'O$, consequently also a point which with $H'T'$ gives an image of FGO . On the other hand, considering the line JK we find points one of which with JK gives an image of FGO , the other of which with $J'K'$ an image of the triangle $F'G'O$. In the case of each of these lines it proves possible to find a point such that an image *either* of FGO or of $F'G'O$ results, but not points such that images of both result.

Fig. 3 *d* now shows the interpretation of these four pairs of images of FGO and $F'G'O$. Each of the four pairs indicates unambiguously one

Fig. 3 (c. d).



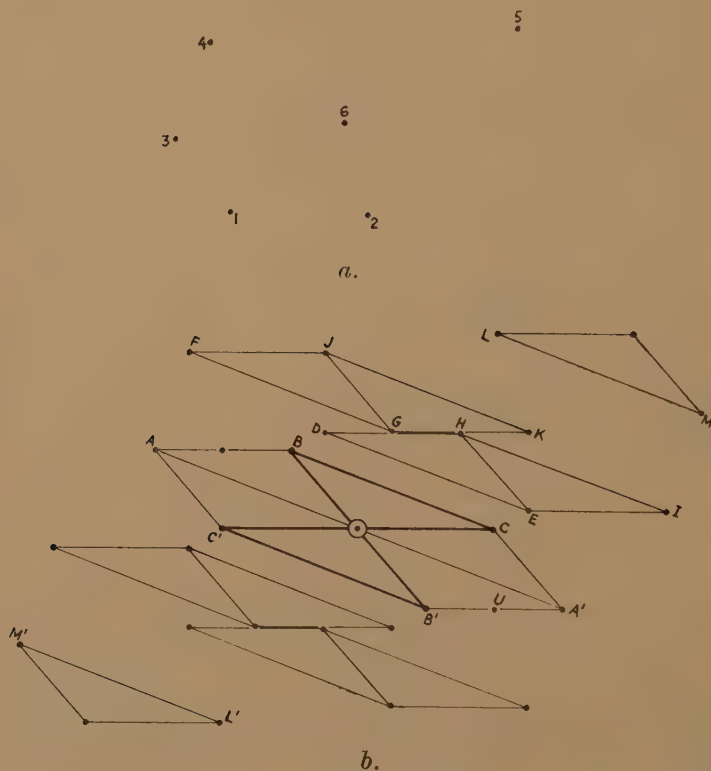
new point in S_1 , and we thus reach a 7-point set. Returning to the original S_2 set, we delete all the points now accounted for and find that the remaining six point pairs represent the interactions of these four new points in S_1 . It follows that this 7-point set is a root of S_2 . On comparing this set (in fig. 2 *d*) with the original S_1 set from which the S_2 set was in fact constructed, we find that the points of the one set are simply the partners of the points in the other set with respect to any point O in their plane.

We notice that the original S_1 set would have resulted if a start had been made by drawing a triangle FGO instead of $F'G'O$. Apart from this choice there has been no possibility of ambiguity at any stage of the argument, and we can therefore assert that the S_2 has as its sole S_1 sets the partner sets shewn in fig. 3 *a* and fig. 3 *d*.

So far the S_1 point sets for the given S_2 point sets have been discussed. If now S_2 be a point intensity set, then the only possible S_1 point sets

have as ground plan the two partner sets already discussed. It is then additionally necessary that giving to the points of these S_1 sets intensities $i_1, i_2, i_3, i_4, i_5, i_6, i_7$, the given S_2 intensities should tally for some values of other of (i_k) with the various $i_m i_p$ intensities. The intensity at the origin in S_2 must be equal to Σi_m^2 and the 21 other intensities must tally with intensities $i_m i_p$ for the appropriate values of m and p . In particular the sum of all the S_2 intensities must be equal to the square

Fig. 4 (a, b).



of the sum of all the S_1 intensities. Here then are 22 equations for seven unknowns. Evidently severe conditions must be imposed on the 22 S_2 intensities if a solution is to be possible. There is always a companion solution to any solution i_1, i_2, \dots, i_7 in the form $-i_1, -i_2, \dots, -i_7$. In this case this is the only multiplicity of intensity values.

The example in fig. 4 a is constructed to show one way in which ambiguity may appear to arise in applying the process we have devised for analysing S_2 point sets. The points in S_1 1, 2, 3, 4, 5, 6 have been so

arranged that $12+34+56=0$, i. e., these three vectors form a closed circuit. This may also be written $34-65=21$. But when we consider this statement in relation to the points in fig. 4 *b* we see that it means that the line from the point 65 to the point 34 is parallel and equal to the line from 1 to 2 in S_1 . Here then is a case in which a parallel displacement of the line 1 to 2 in S_1 into S_2 does not indicate an image of this line 1 to 2 in some third S_1 point.

Considering now the vector map in fig. 4 *b*, we notice there are 15 point pairs, which may indicate a 6-point S_1 . Selecting any point pair, say AA' , we draw all lines in S_2 which are parallel displacements of AO . With six points there can be (excluding AO and OA') only 4 pairs of lines which are images of AO or OA' . But there are in fact six. With the picture of a 6-point S_1 in mind, this can only mean that there are two points in S_1 such that one each of five pairs (including AO , OA') is the image of these points in points of S_1 , and that representatives of two pairs are spurious in the sense that though a parallel displacement of this line joining these points, they are not the image of them in any S_1 point.

We proceed now as before. We select arbitrarily one line, say BC , which directs our attention to the triangle BCO and its partner the triangle $B'C'O$. In an S_1 space we then draw a triangle BCO (fig. 4 *c*), and proceed as in the previous case to find points which, associated with the point pairs under consideration, give images of one or other of these triangles. For one pair of lines, namely LM , $L'M'$, we find one possibility only, and this is interpreted to mean the existence in S_1 of the point Z . With regard to the others there is a choice. Thus we can take JK as spurious and FG as an authentic image of BC , or *vice versa*, and we can take HI as spurious and DE an authentic image of BC , or *vice versa*. The four possibilities are realized in the figs. 4 *c*, *d*, *e*, *f*. Now for any of these four to be satisfactory it is essential that the vector OU (which has not been worked into any of the triangles) should represent the distance between the point Z and one of the two new points, the four possibilities regarding which are shown in figs. 4 *c*, *d*, *e*, *f*. This restricts us to one possibility only, namely, that shown in fig. 4 *d*. As before, the partner of this 6-point set is also a solution of S_2 . The possible source of ambiguity has therefore been dealt with, and in this case also a unique pair of S_1 sets has been found with the given vector map.

We notice that, given any S_1 point set satisfying a given S_2 , we can prove that it (and its partner) constitute the complete solution of the problems in cases where the points in S_1 are such as to exclude the complication illustrated in fig. 4 *a*. In this case there is a closed circuit $34+12+56=0$. For any two points a , b there can always be closed circuits in which ab occurs of the form $ab+bm+ma=0$. If we impose the condition in S_1 that the only closed triangular circuits in which any

two points ab take part are of the form $ab+bm+ma=0$, then this complication is excluded.

The illustrative example pictured in figs. 4*a-f* further shows that even when this condition is not satisfied the number of S_1 sets which share a given S_2 is always severely restricted and ascertainable. The various cases which arise form the subject of a forthcoming publication.

A second complication must be mentioned. In dealing with the 43-point S_2 set shown in fig. 3*b* we found that in no case of the set of five lines CD , EF , etc. was it possible to find two points which would

Fig. 4 (*c-f*).



allow CD (*e. g.*) to form with one a parallel displacement of FGO and with the other a parallel displacement of $F'G'O$. There was therefore no difficulty in interpreting the situation as indicating in the case of each pair of lines one new point in S_1 . Now in figs. 2*a* and 2*b* we considered an S_1 comprising a triangle of points $(1+2+3)$ and a point 4 and its vector map. We imagine for the moment that in the case of the line $(24, 14)$ there exists not only a point, namely 34, which gives with $(24, 14)$ a parallel displacement of the triangle $(23, 13, 0)$, but a second point which with 24 and 14 yields a parallel displacement of $(32, 31, 0)$. But this situation cannot occur unless there is present in S_1 a closed circuit of the

form $24+14+nm=0$. In excluding this possibility we exclude this complication. In a case in which this complication occurs we deal with each of the alternatives in turn, checking up to see if the remaining points in S_2 give satisfactorily the vectors required in each case. The number of alternative S_1 partner sets is evidently easily ascertainable.

It is of course evident that the vast majority of point sets are not vector maps of other sets, just as the vast majority of integers are not the squares of other integers. Of the sets which are vector maps the vast majority will respond to the treatment indicated and have as their S_1 sets one specific set and its partner set. The treatment described has so far limited us to the consideration of S_2 sets which have 3, 6, 10, 15, 21, $\dots \frac{1}{2}n(n-1), \dots$ point pairs. This, of course, though necessary for an S_2 derived from an S_1 which has no identical vectors, is not sufficient to prove for an S_2 that the S_1 sets are of this kind. Thus, if the procedure given above breaks down in the case of an S_2 with $\frac{1}{2}n(n-1)$ point pairs, or if S_2 has a number of point pairs not of this form, then new considerations must be introduced.

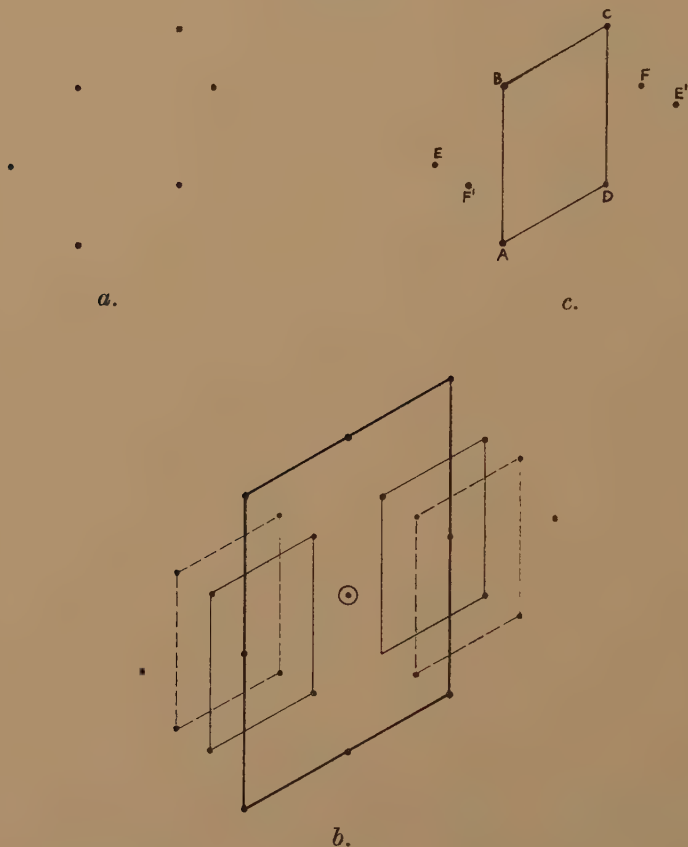
On the present occasion the procedure is to consider the types of degeneracy which arise and to point out how each in turn can be handled. In actual crystallographic practice we are concerned presumably with only S_2 sets which are the vector maps of other sets. Thus sooner or later, by the application of these procedures, one or more S_1 partner sets must emerge. Theoretically, however, it is possible that an S_2 set will not respond to this treatment. In such a case the solution of the problem is that the S_2 in question is not the vector map of any point set.

5. Degenerate Point Sets and their Vector Maps.

To discover the appropriate procedure we study vector maps of "degenerate" sets, *i. e.*, of sets which have coincident vectors, and fig. 5 *a* for example is a case of two pairs of coincident vectors. We notice in fig. 5 *b* a central parallelogram which carries additional points at the midpoints of its sides. These four corner points and the four midpoints comprise with the centre the vector map of a parallelogram of the same orientation and of half its dimensions. This set of points will be called a "mid-loaded" parallelogram, to indicate that there are points at the midpoints of the joins of corner points two at a time. This mid-loaded central parallelogram (the only one in the diagram) suggests that we try the half-size parallelogram ABCD in S_1 shown in fig. 5 *c*. We then find images of ABCD in S_2 which indicate additional points E, F (or their partners in ABCD). No choice except between E and its partner or between F and its partner is encountered in the argument, and we can then assert that these partner sets ABCDEF, CDABE'F' are the only solutions of the problem.

Of great interest also is the case of an S_1 which has two and only two coincident vectors, which therefore must have three collinear points evenly spaced. We can evidently detect the existence of such a situation in S_1 (fig. 6 *a*) by noticing the alignments of points in S_2 (as shown in fig. 6 *b*). To analyse an S_2 such as that shown in fig. 6 *b* we build the three collinear points evenly spaced and then proceed as before to find images of this

Fig. 5.



set in S_2 . Only a restricted and ascertainable number of S_1 partner sets exist in each case.

An example of an S_2 which immediately indicates a considerable amount of degeneracy in its S_1 sets (if any) is shown in fig. 7 *a*. It consists of a central point and nine point pairs trigonally arranged at the corners of a hexagon and at the corners and midpoints of the sides of a hexagon of the same orientation and double the dimensions.

We notice two and only two central "mid-loaded" partner parallelograms, and we accordingly draw, in S_1 spaces, the two corresponding half-size parallelograms ABCD, PQRS (figs. 7 *b* and 7 *c*). Since the number of point pairs is nine, either six or more points are involved in S_1 . We then have to look for two or more additional points in each case. All the parallel displacements of these S_1 parallelograms in S_2 are now remarked. In the first case they indicate two S_1 points E and F or E' and F', or

Fig. 6.

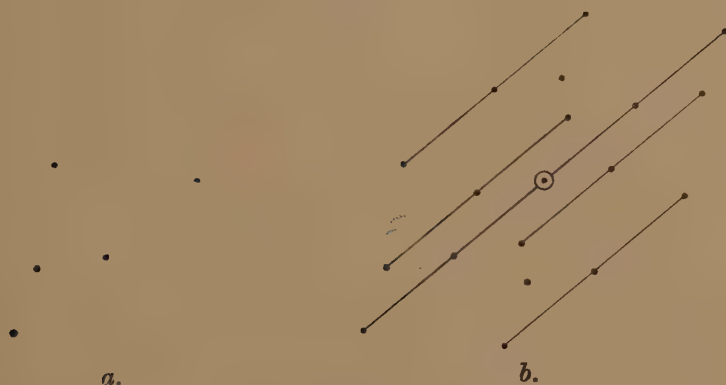
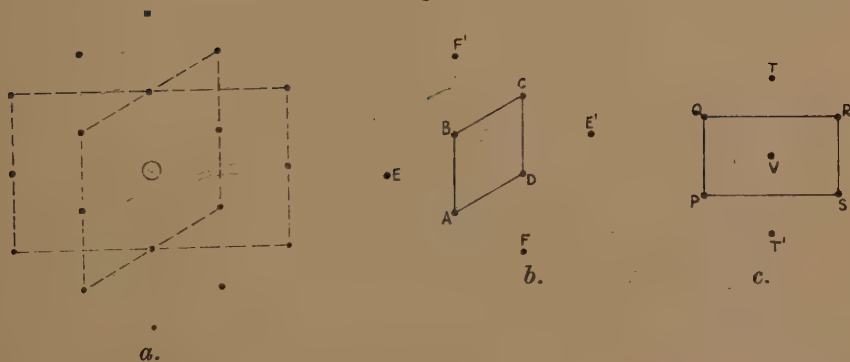


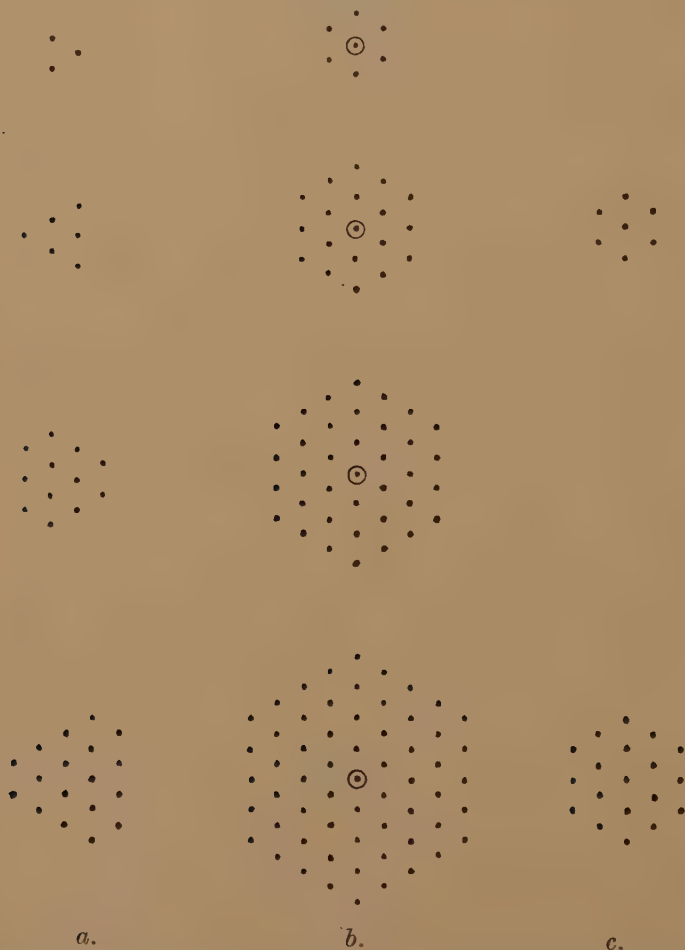
Fig. 7.



E and F' or E' and F, and no other points. In the second case they indicate two points T and T' with or without the point V at the centre of the parallelogram. Nothing further can be got out of this S_2 set, and each of the S_1 sets ABCDEF, ABCDEF' (and their partner sets CDABE'F', CDABE'F), PQRSTT'V (which is its own partner), and PQRSTT' (also its own partner) satisfies S_2 . These then comprise the complete solution of the problem⁽⁴⁾.

In these remarks we are concerned with point sets, not with point intensity sets. In general, though two distinct point sets may have the same vector map, two distinct point set intensities built on these ground plans will have different point intensity vector diagrams.

Fig. 8.



6. *Point Sets belonging to an Equitriangular Mesh System and their Vector Maps.*

This special case directs attention to the general theory of triangular meshes (consisting of equilateral triangles) and its application to vector maps. Evidently the vector map of an S_1 point set will consist wholly of points in such a triangular mesh system when and only when the S_1

To illustrate this point we construct the vector point sets belonging to a series of point sets of this type, as shown in fig. 8 *a*, and append their vector point sets (fig. 8 *b*). These sets all have trigonal symmetry about a point which does not itself lie in the triangular mesh system but is the centre of a single triangle of the system. A second series of sets is also considered (fig. 8 *c*); these are hexagonal, about a point belonging to the system which can therefore also be included. It is of particular

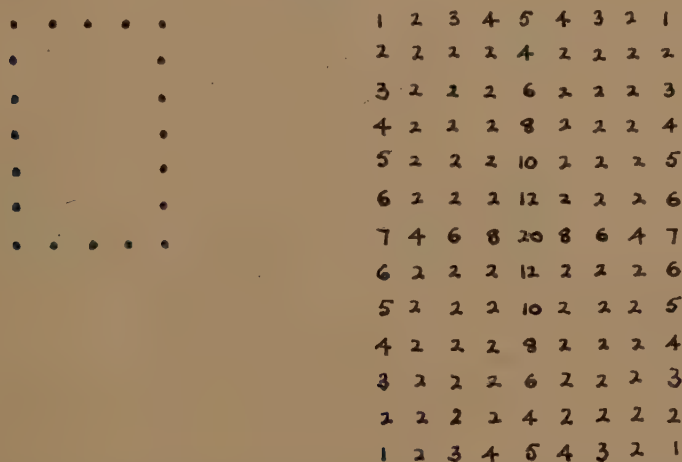
Fig. 10.



importance to notice that the vector point sets of the second series are included among the first set. Thus a system of two, four, . . . concentric hexagons with sides in arithmetic progression, loaded at all points of the triangular mesh to which they belong, have correspondingly two distinct S_1 sets. One of these sets is trigonal, so that any point intensity set built on this plan has either no symmetry or triad symmetry. On the other hand, the second point set is hexagonal. A point intensity set built on this plan can therefore have no symmetry, or dyad, triad, or hexad symmetry.

In this connexion it is also of interest to notice that when point intensity sets are under consideration regions of low and high density can be detected from the vector map direct. This is of special practical importance, since the first application of the new method of solving vector diagrams is to protein crystals, for whose molecules a structure with a low density interior is predicted⁽⁸⁾. As examples we show, in fig. 9, a uniform distribution of intensities on a triangular mesh, say 1, at each point, a distribution of unit intensities at points of a triangular mesh system on a hexagon with zero density within, and a distribution which has a rim two points thick with intensity unity, with zero density within. A careful study of the resulting vector maps shows how we can detect a different density

Fig. 11.



in the interior of a molecule, and how we can assess the thickness of the rim of the molecule.

The theory of a rectangular mesh system can be used for solving vector maps after the manner shown for a triangular mesh. As before an S_1 will have a vector map which belongs to a given rectangular mesh system when, and only when, it itself also belongs to this system. We show examples in fig. 10. The detection of regions of different density is S_1 can also be accomplished direct from the vector maps, as indicated in fig. 11.

7. Point Intensity Sets with a Common Vector Map.

The example discussed on page 101 drew attention to the kind of questions involved when we seek the point intensity sets (if any) corresponding to a given vector point intensity set. In this case it was found

that unless the S_2 intensities satisfy certain conditions there are no S_1 intensity sets, but that when the conditions are satisfied the S_1 intensity sets are confined to the S_1 point sets corresponding to the S_2 point sets with intensities uniquely determined except that a change throughout in the sign of the intensities is permissible. We now draw attention to a highly degenerate case in which a genuine multiplicity of intensity sets (all built on the same ground plan) exists*.

Let us take as the vector map a point at the origin of intensity O and collinear point pairs of intensity A and B, the second pair being twice as far from O as the first pair. This set of five collinear points at equal distances, symmetrically placed with regard to the origin in vector space, has as its unique S_1 set three equally spaced collinear points 1, 2, 3. To consider the solution of the point intensity vector set we give these intensities i_1, i_2, i_3 , and it is then necessary that these intensities should satisfy the equations

$$i_1^2 + i_2^2 + i_3^2 = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$i_1 i_2 + i_2 i_3 = A, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$$i_2 i_3 = B. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Evidently for any satisfactory set (i_1, i_2, i_3) there will be a second satisfactory set ($-i_1, -i_2, -i_3$). To discover any further multiplicity of solutions we may consider the possibility of an alternative set of intensities j_1, j_2, j_3 which must satisfy the equations

$$j_1 + j_2 + j_3 = i_1 + i_2 + i_3, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$j_1 j_2 + j_2 j_3 + j_3 j_1 = i_1 i_2 + i_2 i_3 + i_3 i_1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

$$j_2 j_3 = i_2 i_3. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

In virtue of (4), the equation (5) may be replaced by

$$(j_1 + \omega j_2 + \omega^2 j_3)(i_1 + \omega^2 j_2 + \omega j_3) = (i_1 + \omega i_2 + \omega^2 i_3)(i_1 + \omega^2 i_2 + \omega i_3) \quad . \quad (7)$$

(where ω is a complex cube root of unity), which implies that, for values of α such that equation (6) is satisfied, we have either

$$j_1 + \omega j_2 + \omega^2 j_3 = e^{i\alpha}(i_1 + \omega i_2 + \omega^2 i_3), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$j_1 + \omega^2 j_2 + \omega j_3 = e^{-i\alpha}(i_1 + \omega^2 i_2 + \omega i_3), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

or the similar pair obtained by interchanging i_2 and i_3 . From (4), (8), (9),

$$\left. \begin{aligned} 3j_1 &= (1 + 2 \cos \alpha)i_1 + (1 + 2 \cos (\alpha + 2\pi/3))i_2 + (1 + 2 \cos (\alpha - 2\pi/3))i_3, \\ 3j_2 &= (1 + 2 \cos (\alpha - 2\pi/3))i_1 + (1 + 2 \cos \alpha)i_2 + (1 + 2 \cos (\alpha + 2\pi/3))i_3, \\ 3j_3 &= (1 + 2 \cos (\alpha + 2\pi/3))i_1 + (1 + 2 \cos (\alpha - 2\pi/3))i_2 + (1 + 2 \cos \alpha)i_3 \end{aligned} \right\} \quad (10)$$

* I am indebted to Mr. T. Chaundy for assistance in the solution of this problem.

The equation for α obtained by inserting these values for j_2 and j_3 in equation (6) yields one pair or three pairs of real values for $\cos \alpha$ and $\sin \alpha$, apart from the pair which implies $\alpha=0$ (which joins the original set of intensities).

8. Preliminary Notes on Periodic Point Intensity Sets and their Vector Maps.

It will be remarked that the analysis just given solves a certain problem in the theory of periodic point intensity sets, which are therefore now discussed in a preliminary manner.

A periodic point set is obtained by repeating a finite set or unit A at all points of a lattice L. It has already been shown how the transition from vector maps of finite point sets to vector maps of periodic point sets may be accomplished⁽⁹⁾. The procedure is to select a unit A of the S_1 lattice L, to form its vector map $V(A)$, and to repeat this map at all points of the S_2 lattice L. We then have $L(V(A))$, which is also $V(L(A))$, the vector map of the lattice L.

The simplest case of a one-dimensional lattice comprising n evenly spaced points per cell may first be discussed. Suppose the S_1 set consists then of intensities i_m at evenly spaced points with $i_{m+n}=i_m$. Then if n be odd the vector map in S_2 consists of evenly spaced points with intensity I_m , where

$$\begin{aligned} I_0 &= \Sigma i_m^2, \\ I_1 &= \Sigma i_m i_{m+1} = I_{n-1}, \\ I_2 &= \Sigma i_m i_{m+2} = I_{n-2}, \\ &\dots\dots\dots \end{aligned}$$

the summation being in each over n distinct terms, with, if n be even, the value

$$I_{n/2} = 2 \Sigma i_m i_{m+n/2}.$$

There are therefore $\frac{1}{2}(n+1)$ different intensities in S_2 if n is odd, and $\frac{1}{2}n+1$ if n is even. These will tally with a set of the intensities constructed from n different intensities in S_1 , whatever they may be. But an alternative intensity set (j_m) will also satisfy the problem if their $\frac{1}{2}(n+1)$ —or $(\frac{1}{2}n+1)$ —different functions, shown in the foregoing equations, are equal to the corresponding functions of the set (i_m) . Thus n new intensities (j_m) have to satisfy only $\frac{1}{2}(n+1)$ or $\frac{1}{2}n+1$ equations. A multiply infinite set of (j_m) values is therefore to be expected depending upon the set (i_m) and $\frac{1}{2}(n-1)$ or $\frac{1}{2}n-1$ arbitrary parameters. Thus consider the case when n is three. The equations then are two, namely

$$\begin{aligned} j_1+j_2+j_3 &= i_1+i_2+i_3, \\ j_2j_3+j_2j_1+j_1j_2 &= i_2i_3+i_3i_1+i_1i_2. \end{aligned}$$

To solve them we turn to the analysis given on page 114 and find that for an arbitrary α

$$\begin{aligned}\pm 3j_1 &= (1+2\cos\alpha)i_1 + (1+2\cos(\alpha+2\pi/3))i_2 + (1+2\cos(\alpha-2\pi/3))i_3, \\ \pm 3j_2 &= (1+2\cos(\alpha+2\pi/3))i_1 + (1+2\cos(\alpha-2\pi/3))i_2 + (1+2\cos\alpha)i_3, \\ \pm 3j_3 &= (1+2\cos(\alpha-2\pi/3))i_1 + (1+2\cos\alpha)i_2 + (1+2\cos(\alpha+2\pi/3))i_3.\end{aligned}$$

Correspondingly for the case when n is five the equations are three, and the solution for an arbitrary α and an arbitrary β is

$$\begin{aligned}\pm 5j_1 &= (1+2\cos\alpha+2\cos\beta)i_1 + (1+2\cos(\alpha+2\pi/5)+2\cos(\beta+4\pi/5))i_2, \\ &\quad + (1+2\cos(\alpha+4\pi/5)+2\cos(\beta-2\pi/5))i_3 + (1+2\cos(\alpha-4\pi/5)) \\ &\quad + 2\cos(\beta+2\pi/5)i_4 + (1+2\cos(\alpha-2\pi/5)+2\cos(\beta-2\pi/5))i_5,\end{aligned}$$

with corresponding expressions for j_2, j_3, j_4, j_5 .

For even values of n the procedure is different. Thus for $n=4$ the equations are

$$\begin{aligned}\pm(j_1+j_2+j_3+j_4) &= i_1+i_2+i_3+i_4, \\ (j_1+j_3)(j_2+j_4) &= (i_1+i_3)(i_2+i_4), \\ j_1j_3+j_2j_4 &= i_1i_3+i_2i_4,\end{aligned}$$

and a simply infinite set of solutions may be anticipated. It follows from the equations that, apart from cyclic interchanges and a uniform change of sign,

$$\begin{aligned}j_1+j_3 &= i_1+i_3, \\ j_2+j_4 &= i_2+i_4.\end{aligned}$$

Thus for an arbitrary h we have

$$\begin{aligned}\pm 2j_1, \pm 2j_3 &= i_1+i_3 \pm ((i_1-i_3)^2+4h)^{\frac{1}{2}}, \\ \pm 2j_2, \pm 2j_4 &= i_2+i_4 \pm ((i_2-i_4)^2-4h)^{\frac{1}{2}},\end{aligned}$$

which may also be written in terms of an arbitrary ψ ,

$$\begin{aligned}\pm 2j_1, \pm 2j_3 &= i_1+i_3 \pm i \sin \psi, \\ \pm 2j_2, \pm 2j_4 &= i_2+i_4 \pm i \cos \psi,\end{aligned}$$

where

$$i^2 = (i_1-i_3)^2 + (i_2-i_4)^2.$$

These procedures can be extended to other cases.

As an example of the multiple intensity sets which can share the same vector map we cite the case

$$\dots 1, 2, 4, 1, 2, 4, \dots$$

Other sets with the same vector map, namely

$$\dots 21, 14, 14, 21, 14, 14, \dots,$$

are given by

$$3j_1 = 7 - 4 \cos \alpha - 2\sqrt{3} \sin \alpha,$$

$$3j_2 = 7 - \cos \alpha + 3\sqrt{3} \sin \alpha,$$

$$3j_3 = 7 + 5 \cos \alpha - \sqrt{3} \sin \alpha,$$

which include, in addition to the original set for $\alpha=0$, the set

$$\dots 11/3, 8/3, 2/3, 11/3, 8/3, 2/3, \dots$$

for $\alpha=\pi/3$.

A second example is afforded by the case

$$\dots 3, 7, -2, -5, 3, 7, -2, -5, \dots$$

with the vector map

$$\dots 87, 2, -82, 2, 87, 2, -82, 2, \dots$$

Other sets with the same vector map are given by

$$2j_1, 2j_3 = 1 \pm 13 \sin \psi,$$

$$2j_2, 2j_4 = 2 \pm 13 \cos \psi,$$

which include the sets

$$7, 1, -6, 1, 7, 1, -6, 1, \dots,$$

$$1/2, 15/2, 1/2, -11/2, 1/2, 15/2, 1/2, -11/2, \dots$$

for the values $\psi=\pi/2$ and $\psi=0$ respectively, the original distribution being given by $\psi=\tan^{-1}(5/12)$.

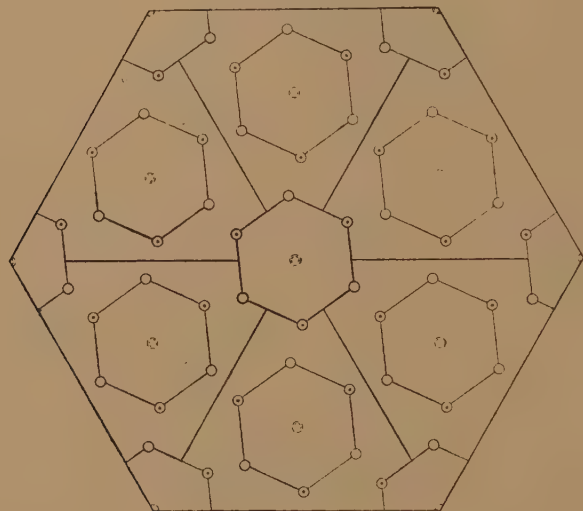
As an example of a two-dimensional S_1 lattice and its vector map in S_2 we take a unit consisting of a hexad of points with a point at its centre (already shown in fig. 8 c and fig. 7 c) and repeat it and its vector map at points of a rhombic lattice as shown in fig. 12 a, b.

The vector map of this unit has already been shown in fig. 7 a, and we know ⁽⁹⁾ that the only corresponding sets are those shown in figs. 7 b and 7 c. It will be noticed that none of the points of the vector maps repeated at the lattice points superpose, and the derivation of two and only two periodic partner sets in S_1 (as shown in figs. 12 a and 12 c) follows.

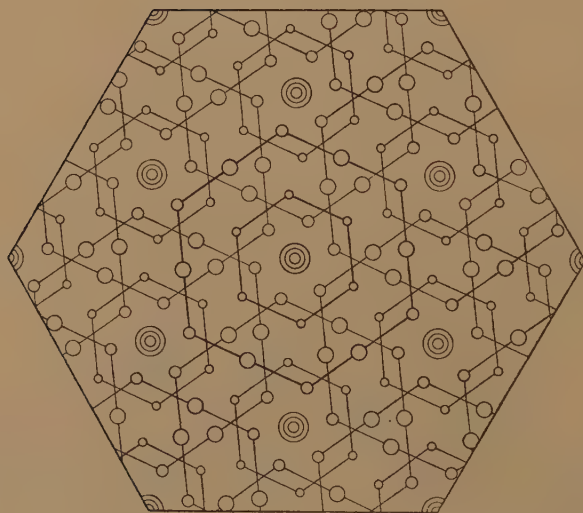
In degenerate cases, in which the lattice parameters are so related to those of the unit that vector maps of different S_1 units have superposed points, complications arise. To make clear the type of complications which can arise, and to indicate how to deal with them, we take the 19-point unit shown in fig. 8 c. We build it into a lattice such that the individual vector maps of these units when placed at all lattice points have some

superposed points. This requires only that the lattice points should belong to the same hexagonal mesh as the points of the S_1 unit

Fig. 12 (*a, b*).



a.



b.

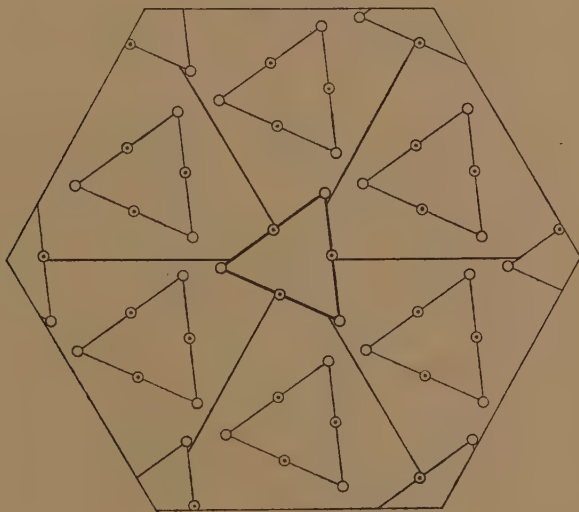
itself, and should not be too far apart. A case of this is shown in fig. 13 *a*. Here all the points in the vector map (fig. 13 *b*) are compound points, *i. e.*, they represent the superposition of two or more points belonging

to vector maps of different units. In fact, the three hexads in S_2 , A, B, C, are compound of order 3, 3, and 4 respectively.

Now suppose that it be required to solve this vector map. We select first a simple unit in this vector map, provided only that it is the vector map of some set of points. In this case we may take the three hexads A_2 , B_3 , C_1 . This is the vector map of the units shown in figs. 7 *b*, 7 *c*, and their partner sets, and of these alone. Hence the only S_1 lattices derived from the vector lattice analyzed into these special A, B, C units are the lattices of these units. One of these is shown in fig. 13 *a*.

These S_1 lattices give the fundamental solution of this vector map, in the sense that all other solutions are derivable from these, for all

Fig. 12 (c).



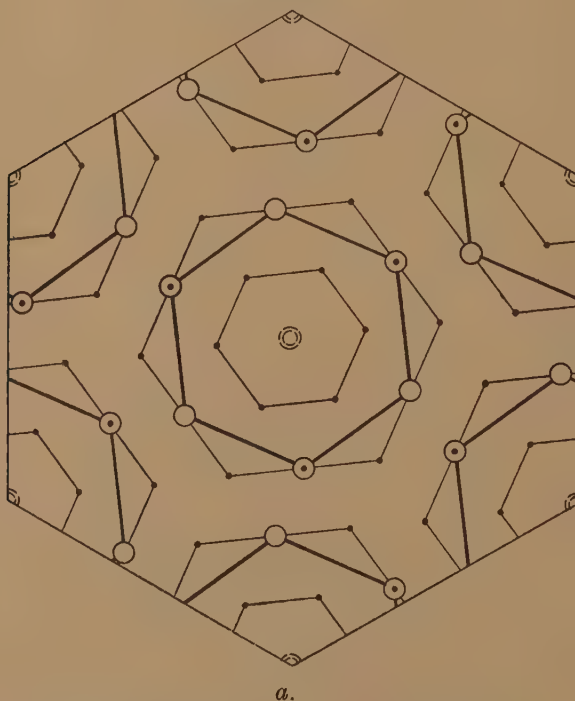
c.

other solutions involving as the vector unit more than one A, one B, and one C hexad require that one or more of the points A, B, C be compound. We take any one fundamental solution, namely that shown in fig. 13 *a*; we now try introducing points into the S_1 unit such that some or all of the inter-unit vectors are now also intra-unit vectors and such that no new vectors are introduced. This is possible only in degenerate cases, being due to the existence of closed circuits made up of inter- and intra-unit vectors. In the present case we are restricted to two new hexads. This also constitutes a solution of the problem in view of the fact that all the new vectors are already present as inter-unit or intra-unit vectors. In fact this is the actual S_1 lattice from which the vector map in question was constructed. It represents an S_1 unit which has as

its vector map a share of three different A hexads, of three different B hexads, and of four different C hexads.

This procedure deals successfully with all degenerate cases of this kind. The degeneracy of course consists in the fact that the S_1 unit belonging to a certain hexagonal or equitriangular mesh system was so placed in the lattice that all the S_1 units belong to the equitriangular mesh system defined by a single unit. Given any such mesh system, then the vector map of an S_1 lattice, all of whose points belong to this system, will also

Fig. 13 (a).

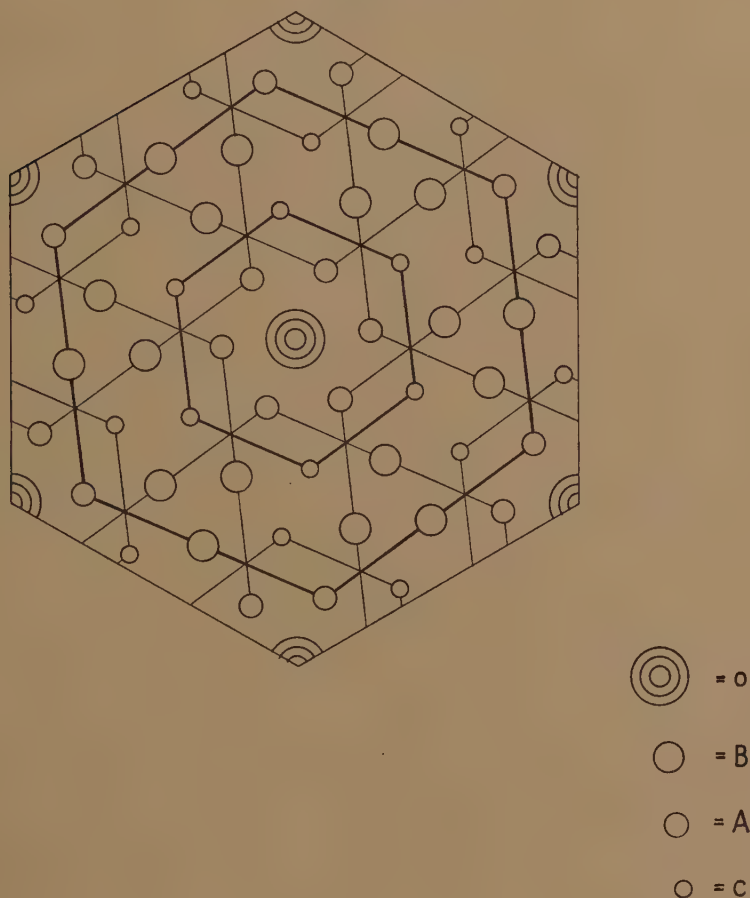


belong to the same system. It is therefore not difficult to construct such degenerate cases nor to find the complete solution of vector maps showing this type of degeneracy.

We would emphasize that, in practical cases in which there is known to be one molecule per cell in the crystal lattice, the principles introduced above allow a systematic approach to be made to the problem of the elucidation not only of the general crystal structure, but also of molecular structure. Thus, in the case considered above the fundamental solution gives a 6-point unit for the S_1 lattice, *i. e.*, a molecule containing six points, or a 7-point unit, *i. e.*, a molecule containing seven points (see fig. 13 b).

Considering either of these solutions, we can derive from it the additional solution shown in fig. 13 *a*, which gives a 19-point molecule. In the case of all derivative solutions it should be noticed that the supposition is necessarily introduced that there happen to be in

Fig. 13 (*b*).



b.

the molecule points whose vector distances are identical with vector distances between points in two different molecules. Without this supposition we are restricted to the fundamental solutions only.

In practical cases we would in all cases first submit for consideration only the fundamental solutions. Only in cases where there is reason to discard the fundamental solutions should derived solutions be considered,

since these on the one hand unnecessarily introduce additional points, and, on the other, necessarily introduce a gratuitous assumption of congruence between inter-molecular and intra-molecular vectors.

I am indebted to Prof. E. H. Neville for help in the preparation of this paper.

References.

- (1) Wrinch, J. Amer. Chem. Soc. lx. p. 2005 (1938).
- (2) Wrinch and Langmuir, J. Amer. Chem. Soc. lx. p. 2247 (1938).
- (3) Langmuir and Wrinch, 'Nature,' cxlii. p. 581 (1938).
- (4) Wrinch, 'Nature,' cxlii. p. 956 (1938).
- (5) Patterson, Phys. Rev. xlv. p. 372 (1934); *Z. Krist.* xc. p. 517 (1935).
- (6) Neville, 'Nature,' cxlii. p. 994 (1938).
- (7) See, e. g., Crowfoot, Proc. Roy. Soc. A, clxiv. p. 580 (1938).
- (8) Wrinch, 'Science,' lxxxv. p. 566 (1937); Proc. Roy. Soc. A, clxi. p. 505 (1937); Trans. Far. Soc. xxxiii. p. 1368 (1937).
- (9) Wrinch (in course of publication).

XI. Symmetrical Figures on Circular Plates and Membranes.

By R. C. COLWELL, Ph.D., J. K. STEWART, Ph.D., and
A. W. FRIEND, M.S., West Virginia University *.

[Received July 22, 1938.]

[Plate II.]

FROM the time that Kirchhoff developed his solution for a circular plate, it has been supposed that the two equations

$$\{J_n(kr) + \lambda J_n(ikr)\} \cos n(\theta - \alpha_n) = 0 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and

$$J_n(kr) \cos n(\theta - \alpha_n) = 0 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

give all the solutions for a plate or membrane. They represent circles and diameters. However, in certain cases deviations occur which are usually attributed to irregularities in the plate. We shall show that the symmetrical patterns other than circles and diameters are not due solely to irregularities, but arise from the fact that equation (2) is a particular and not a general solution. A more general solution may be deduced by adding together two particular solutions. This process is exactly analogous to the method used with square plates †, but is subject to certain limitations. Incidentally we use equation (2) rather than equation (1) because for higher frequencies λ approaches zero and the equations are identical. This means physically that for every figure on a circular plate, a corresponding one appears upon a circular membrane. The two figures will look alike, although the spacing on the plate will differ from the spacing on the membrane. This, too, is analogous to the fact that the figures on square plates involve cosine functions, while those on membranes require sine functions.

When ka is great,

$$ka = \frac{\pi}{2}(n + 2h), \quad . \quad . \quad . \quad . \quad . \quad (3)$$

in which n (an integer) gives the total number of diameters and h (an integer) gives the total number of circles ‡. The pitch is approximately unaltered when any number is subtracted from h provided twice that number be added to n . In raising the pitch nodal circles have twice

* Communicated by the Authors.

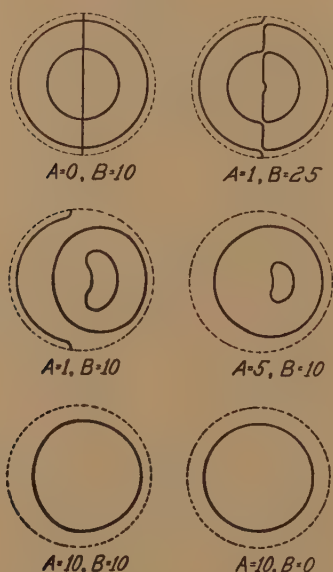
† Rayleigh, 'Sound,' vol. i. Chapters IX. and X.

‡ Rayleigh, *ibid.* p. 361.

the effect of nodal diameters. The physical reason for this phenomenon is that all diameters are really parts of hyperbolas. Hence, the removal of two diameters is equivalent to the removal of one hyperbola. So that the rule is "remove *one hyperbolic curve* from the diameters, replace it with *one circle* (or ellipse) on the plate." Since two different patterns will appear on the plate for the same frequency, it is legitimate to write a second solution in the form

$$J_m(kr) \cos m(\theta - \alpha_m) = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Fig. 1.



provided the m and the n are compatible with equation (4). This is combined with equation (2) to form the terms

$$AJ_n(kr) \cos n(\theta - \alpha_n) - BJ_m(kr) \cos m(\theta - \alpha_m) = 0. \quad . \quad . \quad . \quad (5)$$

To construct the graph of such an equation let us reduce it to the form

$$y = J_m(kr), \quad y = KJ_n(kr), \quad K = \frac{A \cos n(\theta - \alpha_n)}{B \cos m(\theta - \alpha_m)}. \quad . \quad . \quad . \quad (6)$$

It will be understood that equation (6) will give rise to many beautiful mathematical figures when m and n are given different values; but only those patterns which satisfy equation (4) will appear upon the experimental plates for a pure tone. On account of the fact that overtones

may also be present in the notes of the vibrating plate, figures will be seen which are not limited to the relations of equation (4). A new general equation therefore appears in the form

$$y=J_m(k'r), \quad y=KJ_n(kr), \quad K=\frac{A \cos n(\theta-\alpha_n)}{B \cos m(\theta-\alpha_m)}. \quad (7)$$

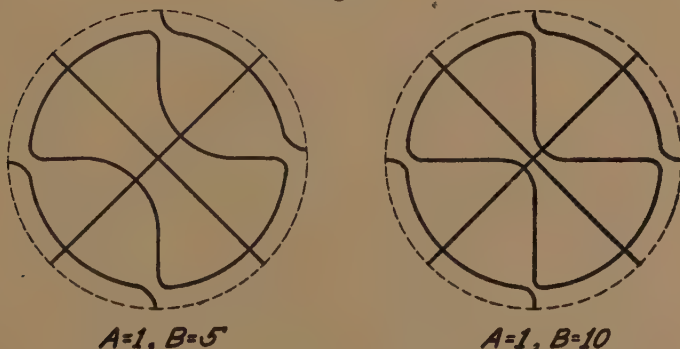
The general equation (7) will be used in the calculations. For example, let the curve be intermediate between two circles with one diameter and one circle. The Kirchhoff equations give

$$AJ_0\left(\frac{3\cdot014r}{a}\right)-BJ_1\left(\frac{7\cdot737r}{a}\right)\cos\theta=0. \quad (8)$$

This equation is plotted for different values of A and B (fig. 1) showing how the Chladni pattern gradually changes from one form to the other. Fig. 2 is a plot of the equation

$$AJ_2\left(\frac{5\cdot937r}{a}\right)\cos 2\theta+\frac{B}{2}J_4\left(\frac{8\cdot55r}{a}\right)\cos 4\left(\theta+\frac{\pi}{8}\right)=0. \quad (9)$$

Fig. 2.



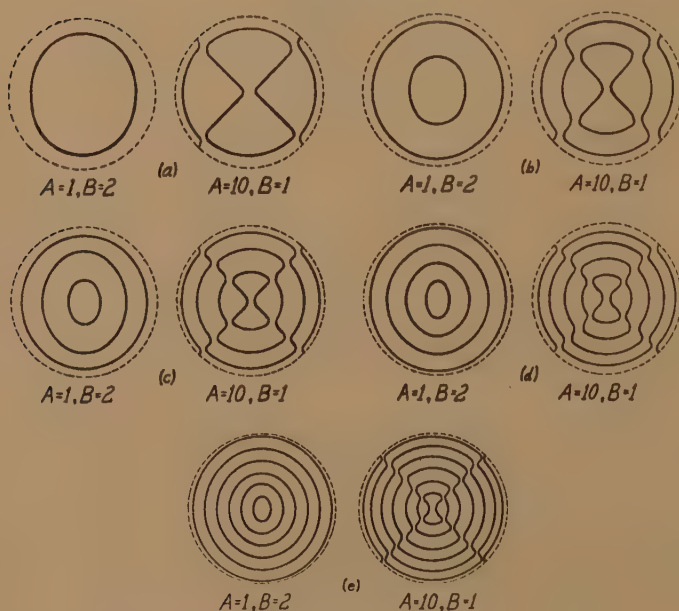
The graphs in fig. 3 show how the nodal lines are moved about by variations in A and B.

$$\left. \begin{aligned} AJ_2\left(\frac{5\cdot937r}{a}\right)\cos 2\theta+BJ_0\left(\frac{3\cdot014r}{a}\right) &=0. \quad (\text{Fig. } 3a.) \\ AJ_2\left(\frac{10\cdot55r}{a}\right)\cos 2\theta+BJ_0\left(\frac{6\cdot209r}{a}\right) &=0. \quad (\text{Fig. } 3b.) \\ AJ_2\left(\frac{13\cdot86r}{a}\right)\cos 2\theta+BJ_0\left(\frac{9\cdot370r}{a}\right) &=0. \quad (\text{Fig. } 3c.) \\ AJ_2\left(\frac{15\cdot58r}{a}\right)\cos 2\theta+BJ_0\left(\frac{12\cdot53r}{a}\right) &=0. \quad (\text{Fig. } 3d.) \\ AJ_2\left(\frac{21\cdot117r}{a}\right)\cos 2\theta+BJ_0\left(\frac{18\cdot071r}{a}\right) &=0. \quad (\text{Fig. } 3e.) \end{aligned} \right\} \quad (10)$$

Fig. 4 is shown because we have been able to find one member of this family experimentally. The equations are for $A=10$, $B=1$.

$$\left. \begin{aligned} AJ_3\left(\frac{7.274r}{a}\right) \cos 3\theta + BJ_0\left(\frac{3.614r}{a}\right) &= 0. \\ AJ_3\left(\frac{10.55r}{a}\right) \cos 3\theta + BJ_0\left(\frac{6.209r}{a}\right) &= 0. \\ AJ_3\left(\frac{13.86r}{a}\right) \cos 3\theta + BJ_0\left(\frac{9.37r}{a}\right) &= 0. \\ AJ_3\left(\frac{17.05r}{a}\right) \cos 3\theta + BJ_0\left(\frac{12.53r}{a}\right) &= 0. \end{aligned} \right\} \dots \dots (11)$$

Fig. 3.



The plot of the equation

$$J_2\left(\frac{5.937r}{a}\right) \cos 2\theta + 10J_6\left(\frac{11.00r}{a}\right) \cos 6\theta = 0 \quad \dots \dots (12)$$

appears in fig. 5 *a*. It is inserted because it resembles a certain pattern which has been obtained experimentally.

The pattern shown in fig. 5 *b* has also been obtained on the plates, and it possesses a type of vibration which has not been discussed heretofore. Since each ring of a circular plate can vibrate between the roots of the Bessel function defining the motion, flat rings (washers) of various com-

plexities can be expected together with diameters. Thus, in fig. 5 *b*, the inner rings are bounded by circles and have no other patterns upon them. The outer ring, however, is vibrating in segments independently of the others. In fig. 6 *a* suppose that the roots of the Bessel function are ρ and $\rho+a$ with $\rho \gg a$. Then the portion of the ring between $\theta=0$ and $\theta=2\pi/c$ (where c is the number of similar segments in the vibrating

Fig. 4.

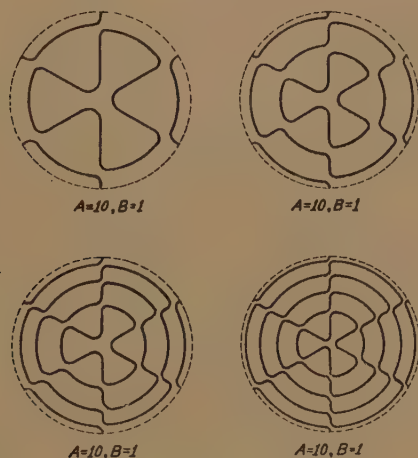
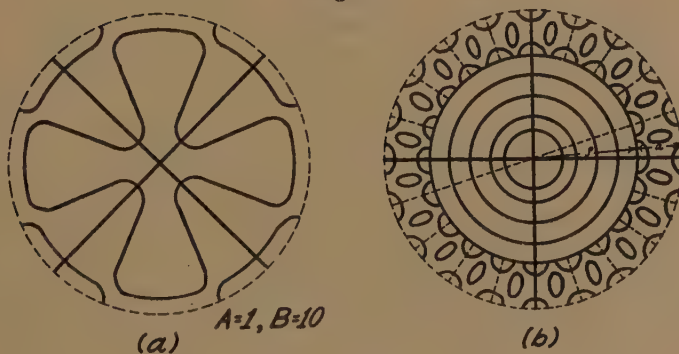


Fig. 5.



ring) is approximately a square, provided a is nearly equal to b . We may, therefore, expect nodal designs similar to those on square plates repeated c times in this ring. These designs are not necessarily separated by nodal lines, although this can happen.

Beginning with the usual equation for square plates,

$$W = A \cos \frac{m\pi x}{a} \cos \frac{n\pi y}{a} + B \cos \frac{n\pi x}{a} \cos \frac{m\pi y}{a} = 0, \quad . \quad . \quad . \quad (13)$$

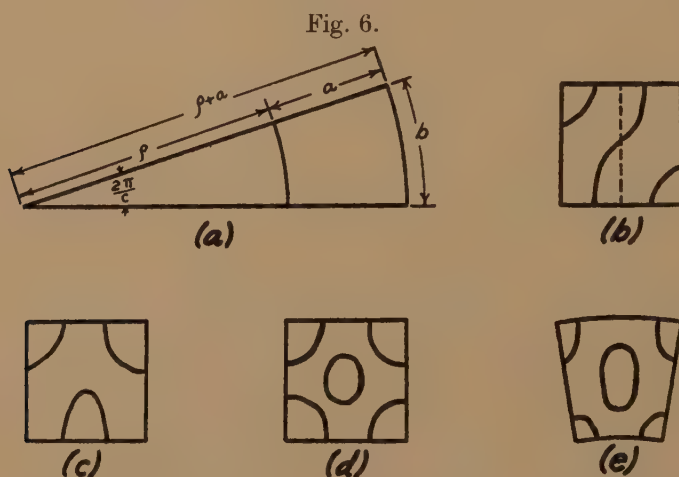
we get by trial the equation of fig. 6 *b* in the form

$$W = 5 \cos \frac{\pi x}{a} \cos \frac{2\pi y}{a} + \cos \frac{2\pi x}{a} \cos \frac{\pi y}{a} = 0. \quad . \quad . \quad . \quad (14)$$

Cutting this figure in two by a vertical medial line and turning one half upside down, we obtain fig. 6 *c*. Doubling the figure, we obtain fig. 6 *d*. This is transformed into fig. 6 *e* by using the general formula

$$W = A \cos \frac{m\pi\rho}{a} \cos \frac{n\pi\theta}{a} + B \cos \frac{n\pi\rho}{a} \cos \frac{m\pi\theta}{a} = 0, \quad . \quad . \quad . \quad (15)$$

in which $x = \rho$ and $y = \rho\theta$.



Twenty of these wedge-shaped segments fitted together will form the outer ring of fig. 5 *b*. In this way, many new figures may be calculated mathematically and then found experimentally with the electronic oscillator. The numerical values used in the Bessel functions will be found in a previous article *. The effect of two simultaneous normal components upon a circular plate has also been discussed briefly by Rayleigh and at greater length by M. D. Waller †.

* R. C. Colwell and H. C. Hardy, *Phil. Mag.* ser. 7, xxiv, p. 1041.

† M. D. Waller, *Proc. Phys. Soc.* 50, p. 76.

XII. Notices respecting New Books.

Traité du Calcul des Probabilités et de ses Applications. Tome i. fascicule 3: *Recherches Théoriques Modernes.* By M. FRÉCHET. Premier livre, 1937. [Pp. 308. Price 125 fr.] Second livre, 1938. [Pp. 315. Price 130 fr.] (Paris: Gauthier-Villars.)

THE great modern encyclopedia of probability, commenced by Borel in 1925, is now approaching completion, and the end has even been promised for this year (1938). But already it has burst its bounds and expanded into a new series of monographs, of which two have been issued and others are in preparation. These have taken the place of what was to have been tome iv. fascicule 3, 'Compléments Divers,' and are apparently designed to carry forward any part of the encyclopedia which is ripe for extension. The first monograph, by P. Lévy, springs indeed from the growing point represented by the present two volumes, and there is, consequently, some repetition of introductory ideas. But the field of modern developments in the "pure" theory of probability is so wide and is so little known to those who might be interested, that the excellent survey made here by Fréchet of the parts with which he has been principally concerned is very welcome. It has, in fact, been very difficult to keep in touch with the great advances made, both in the underlying principles and in the main problems, since the encyclopedia began; much of the published work has been in the journals of Eastern Europe, and some of the progress made by Russian mathematicians in particular takes a prominent place in the books by Cramér and Uspensky which were recently reviewed in this Journal.

Book i. deals with definitions and the fundamental theory of random variables. After a detailed discussion of moments and characteristic functions there is a comprehensive account of the Bienaymé-Tchebicchef inequality and all its extensions. Then the various types of convergence of series of random variables are dealt with, the many points of view being woven into a continuous account. Those whose acquaintance with the theory has been limited to the classical ideas and problems will find here a lucid exposition of the change in point of view—a change which can be roughly stated as the transition from the "arithmetic and algebra" to the "analysis" of the subject. The power of modern analysis to work with very general assumptions, the use of abstract spaces as a mode of dealing with functional relations of the most general type, are shown in such a way that previous familiarity with the ideas is not essential to a grasp of the results achieved.

In Book ii. the main topic is the theory of simple Markoff chains of events, but its relation to Poincaré's concept of arbitrary functions is made plain and the way is prepared for further extensions. Here again the generality of the treatment, which links into a satisfying whole numerous tentative applications of the classical theory to "entangled" probabilities, is preserved from vagueness by detailed application to special problems such as the shuffling of a pack of cards. In each book there is a detailed table of contents, which is probably more useful than an index would be, and a bibliography, that in Book ii. being a continuation of the one given by Hostinsky—'Mémorial des Sciences Mathématiques, no. 52 (1931).'

Some Notes on Least Squares. By W. EDWARDS DEMING, Ph.D. [Pp. 181.] (The Graduate School, Department of Agriculture, Washington, 1938. Price \$1.50.) (Mimeographed.)

As explained in the preface, this is not offered as a text-book on theory of errors or even on least squares as a general process, but as part of a course on analysis of observations given at the Graduate School under the United States Department of Agriculture. It is reproduced from excellent typed copy and strongly bound in a stiff paper cover, but, unfortunately, not so as to lie open conveniently. The main part of the text is based on the author's papers in this Journal a few years ago and assumes a slight acquaintance with modern statistical tests of significance. This could be obtained, for instance, from the paper by Deming and Birge in the 'Reviews of Modern Physics, 1934,' which has been issued as a pamphlet obtainable from the School.

Considered as a supplement to theoretical courses the book has considerable merits. The treatment is consistently based on weighting and examination of standard errors in all variables, it is thorough and clear in arrangement, it covers both the geometrical problems such as are met with in surveying and the general problem of curve fitting. Well selected references are made to other sources for the theory, and the assumption of normality involved in the tests is emphasized now and again. Three particular problems not always properly dealt with are carefully and soundly examined, viz., (i.) the effect on weighting of changing the form of the function fitted, e. g., from $y=ae^{bx}$ to $\log y=\alpha+\beta x$, (ii.) the indeterminacy which often arises from unsuitable data, (iii.) the relation of systematic errors to the estimation of standard errors.

On the other hand, the number of worked numerical examples is very small, and only one scheme for solution of the normal equations is discussed in full. Bare mention is made of the use of orthogonal polynomials, and the statistical developments in correlation work generally are regarded as outside the scope of the book.

An Introduction to Laboratory Technique. By A. J. ANSLEY. [Pp. 313+121 figs.] (MacMillan. Price 12s. 6d.)

THIS book, with a foreword by Professor F. H. Newman, is one which will surely commend itself to a wide circle of people. Teachers, laboratory workers, students young and old, will all find most valuable information and many useful "tips" in its pages. The writer of the book has himself been a laboratory assistant for many years and knows what he is talking about, and the manner of presentation, matter, and literary style are not only suitable, but definitely good.

In this laboratory encyclopedia there are fourteen chapters dealing with almost everything that can possibly be wanted in an ordinary laboratory. The care of equipment, cements, electrical instruments and motors, electroplating, glass blowing, lantern-slides, production of mirrors, soldering, storage cells—these are only some of the matters with which the author deals. In many cases references to descriptive papers are given—but it is only a very minor criticism to suggest that in an essentially practical book of this sort additional value would be conferred on it if the sources of supply of some of the obviously valuable materials mentioned could be made available.

It would be easy to draw up a list of omissions—not a very long one!—but the surprising thing is that Mr. Ansley has contrived to put in so much. The book can be thoroughly recommended.

R. W.

Summable Series and Convergence Factors. By C. N. MOORE. [Pp. vi+105.] (American Math. Soc. Colloquium Publications, vol. xxii. Price \$2.)

CÉSARO has given a whole class of methods of summation, of which summation by arithmetic means is the simplest example, and NÖRLAND has given a general scheme of methods of summation into which all Césaro methods fit. If we single out any one of these methods of summation, we speak of convergence factors for summable series, meaning factors f_n which make $\sum u_n f_n$ converge whenever $\sum u_n$ is summable by the method selected.

The present work is devoted to the pure theory of convergence factors, i. e., it gives the conditions that a set of factors f_n should be convergence factors for summable series. It gives an admirable account of this theory, the subject-matter being divided up, classified, and ordered in such a way as to give an excellent bird's-eye view of the whole. In each stage of the subject simple series are first considered, then double series, and, finally, multiple series in general. In the last chapter the whole theory for double and multiple series is recast, using the idea of restricted convergence in place of convergence. There are also a few pages on "completely convergent" double series. Throughout the book there is a parallel development of convergence factors of Type I. and those of Type II., the former being factors which merely produce convergence, the latter being factors which, like x^n , depend on a parameter, and all approach 1 as the parameter approaches some limiting value.

The book is in the somewhat severe style usual in American books on pure mathematics. This is not in itself a fault, for such a style has many advantages, but associated with the style is the almost complete absence of applications of the theory. From the title one might have expected a discussion of the relations between different methods of summation. This could have been given for Césaro's methods as an application of the theory: it would have clothed the skeleton in a little flesh. Such phrases as "however small" and "powerful method of summation" are conspicuously absent from this book.

Engineering Electronics. By DONALD G. FINK. [Pp. 358+134 figs.] (McGraw-Hill Publishing Company, Ltd. Price 21s.)

WRITTEN for engineers with a foundation and background of electrical knowledge but no specific training in electron ideas the book, in fact, is based on a series of lectures given to such men from the Westinghouse Lamp Company of America. The preface points out that this lecturing experience showed the necessity of providing:

- (1) An understanding of electronic conduction,
- (2) A familiarity with different types of "electron tubes,"
- (3) Knowledge of circuits in which such tubes are used.

There are consequently three sections in the book on Physical Electronics, Electron Tubes, Electron-tube Applications—or, as we should say in this country, Gaseous Conduction, Valves, and Valve Circuits.

A short introductory chapter gives picturesque illustrations of the general subject-matter. An example of this "local colour" is the reference to the repeated amplification in the New York-San Francisco telephone-line, in which it is stated that fifteen successive stages of amplification give a total voice energy amplification of 1,000,000,000,000 times.

The book is naturally elementary in character, but very suitable for the purpose in view. Clear ideas are given of the apparatus described, and the book is lavishly sprinkled with diagrams, charts, photographs, and simple formulæ.

The original intention seems to be admirably fulfilled—as a semi-popular account appealing to beginners in the subject the book may also recommend itself. R. W.

Physik für Studieren an Technischen Hochschulen und Universitäten. By Ing. Dr. P. WESSEL. [Pp. 514, 277 diagrams.] (Published by Ernst Reinhardt, München. Price 4.90 R.M.)

THIS brave book attempts to present in some four hundred pages an introductory survey of Physics, starting with fundamental principles of mechanics and leading up to recent ideas on the structure of atoms, relativity, and quantum mechanics. It is inevitable that such an ambitious attempt in so short a space should be not wholly successful. But the measure of success attained is quite astonishing. Concentrating on essential principles, the author has been able to combine clarity with brevity to a most unusual degree. The presentation is greatly helped by diagrams which are not only clear but in many cases most original in the way in which the main idea is emphasized. While, no doubt, teachers will hesitate to recommend to students a textbook which devotes only five or six pages to Sound and another twenty-five to Heat, while dealing with such topics as resolving power, oscillating circuits, and nuclear structure, many will welcome the freshness of approach and few will fail to gain something from its perusal.

There is a catechism of 1447 short questions and a number of useful tables.

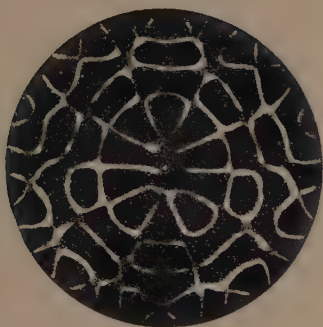
Fluorescence and Phosphorescence. By E. HIRSCHLAFF. Methuen's Monographs on Physical Subjects. [Pp. 130.] (Methuen and Company, Ltd. Price 3s. 6d.)

THIS little book has the advantage of being the only account in English of some of the modern work on luminescence, and will therefore be welcomed. It can, however, hardly be said to fulfil the intention of the series—to give a compact statement of the modern position in each subject. It is unfortunate that publication was not delayed until after the recent conference of the Faraday Society on Luminescence. The author would then have been better able to acquaint himself with the large amount of theoretical and experimental investigations carried out in the last few years, and to have given a more balanced survey of the whole field. No reference is made, for instance, to the recent applications of the zone theory of crystals to the luminescence of solids. A somewhat irritating feature is the frequent citation of results without adequate discussion or even reference.

[The Editors do not hold themselves responsible for the views expressed by their correspondents.]



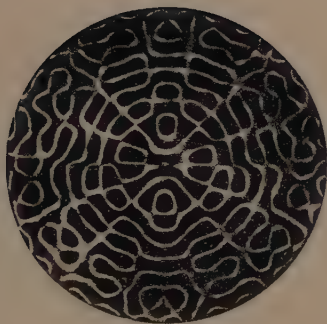
O . POSITION OF UNDEFLECTED SPOT.
FOR PHOTOGRAPH 6



(a)



(b)



(c)



(d)



(e)



(f)

XIII. *The Force-time Law governing the Impact
of a Hammer on a Stretched String.*

By N. DAVY, J. H. LITTLEWOOD, and M. McCAIG,
University College, Nottingham *.

[Received October 7, 1938.]

[Plate III.]

SUMMARY.

AN attempt is made to measure the instantaneous force of impact of a hammer on a string by a piezoelectric method, at any moment while the two are in contact. Cathode ray oscillograms, equivalent to force-time diagrams, are exhibited for both soft and hard hammers.

Statement of the Problem.

MANY papers on the struck string, both experimental and theoretical, have been published, but, as far as can be ascertained, no formal experimental investigation of the law connecting the force between hammer and string with the time, during the period of contact, has appeared. The practice of authors requiring such a law in connexion with their theories has been either to assume a form of the law *a priori* or to make certain assumptions and thence to deduce a law theoretically. The method of the present paper would seem to afford an experimental means of obtaining the force-time law of impact, at any rate graphically. Harmonic analysis can then be used to obtain an algebraic form of the law.

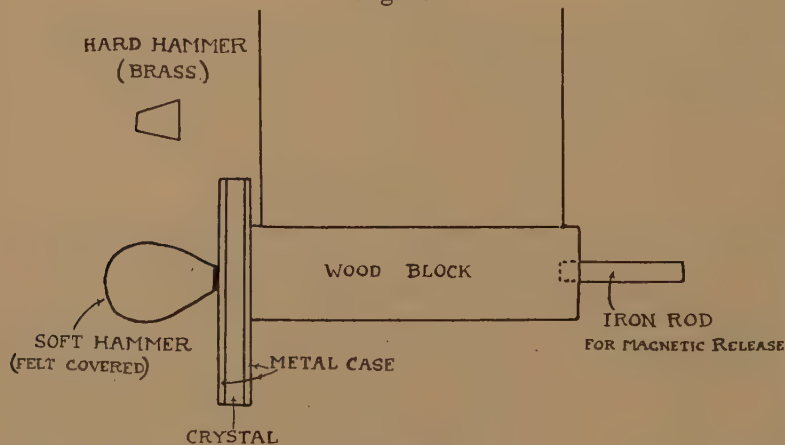
Method of Investigation and Theory.

The general plan of the work was to sandwich a piezoelectric crystal in the head of a hammer, to let the hammer strike a string so that piezoelectric potential differences were produced between metal plates on opposite sides of the crystal, to amplify these potential differences, and by connecting to a cathode ray oscillograph provided with a linear time base, to obtain an oscillogram whose abscissæ were proportional to time and whose ordinates were functions of the instantaneous force between hammer and string. The following theory combined with the experimental results indicates that in the oscillograms obtained the ordinates

* Communicated by Prof. L. F. Bates, Ph.D., D.Sc.

were approximately directly proportional to the instantaneous forces. A hammer in the form of the bob of a pendulum was used (see fig. 1). The bob was supported by four thin wires, two of which also acted as leads connecting the crystal to the amplifier. On breaking the current of an electromagnet the displaced bob was released, and after falling down a circular path struck the wire when the centre of gravity reached its lowest point. Knowing the position of the starting-point and of the point to which the centre of gravity of the bob rebounded, and neglecting the effect of air resistance, the momenta just before and just after impact could be calculated. The sum of these gave the change of momentum during impact, which lasted for a finite time of a few hundredths of a second. The impact caused a corresponding impact figure to appear

Fig. 1.



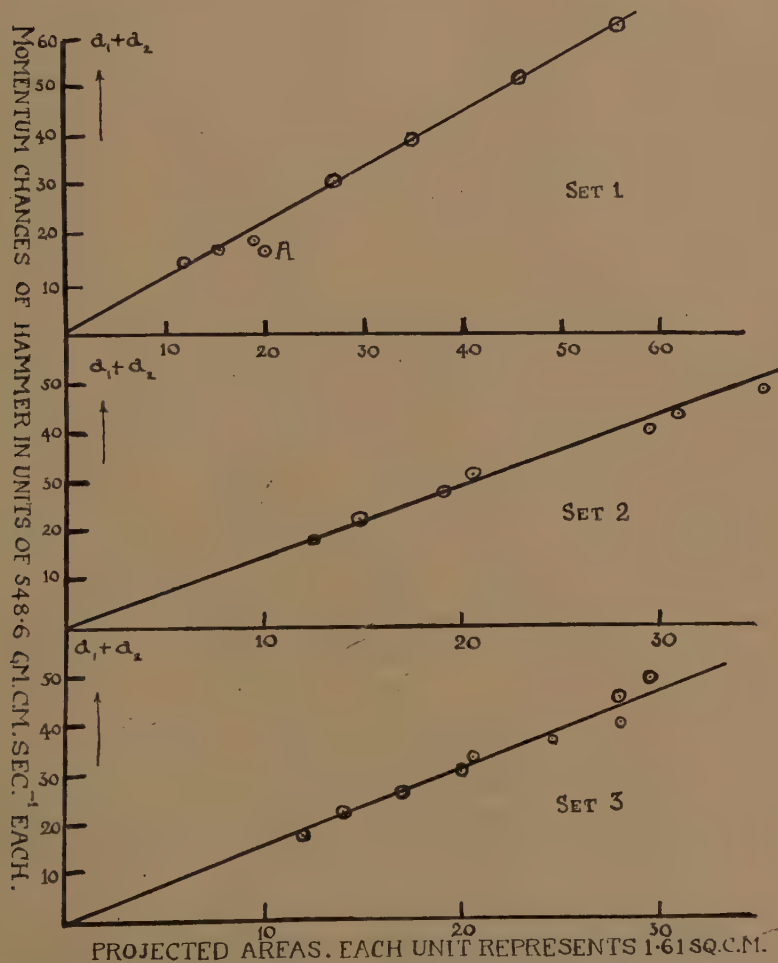
The hammer.

on the screen of the oscillograph. Having arranged for these oscillograms to be formed in the centre of the screen, a set of photographs was taken with various values of change of momentum. Under constant magnification the photographed oscillograms were projected on squared paper, and in each case the area between the curve and the time axis was measured. A graph was plotted with areas as abscissæ and corresponding changes of momentum as ordinates. In several sets of experiments the graph appeared to be a straight line passing through or very nearly through the origin (see fig. 2). Calculated by Awbery's method, the intercepts on the axes were very small. Hence the areas of the projected oscillograms were, at any rate approximately, directly proportional to the corresponding changes of momenta.

The assumption that the ordinates of the projected oscillograms are proportional to the instantaneous forces leads to the same result. For,

let the coordinates of any point on a projected oscillogram be x, y . By arrangement a linear time base gives $x=k_1t$, where k_1 is a constant and t is the time measured from the beginning of impact. Assume that $y=k_2F$ where k_2 is a constant and F is the instantaneous force

Fig. 2.



Graphs connecting projected areas of oscillograms and corresponding momentum changes of the hammer.

exerted by the string on the hammer. Now the area of the projected oscillogram is $\int y \cdot dx$, the integral being taken over the whole duration of impact. This integral is equal to $k_1 k_2 \int F \cdot dt$. But $\int F \cdot dt$ is the change of momentum of the hammer and $k_1 k_2$ is a constant, so that with the

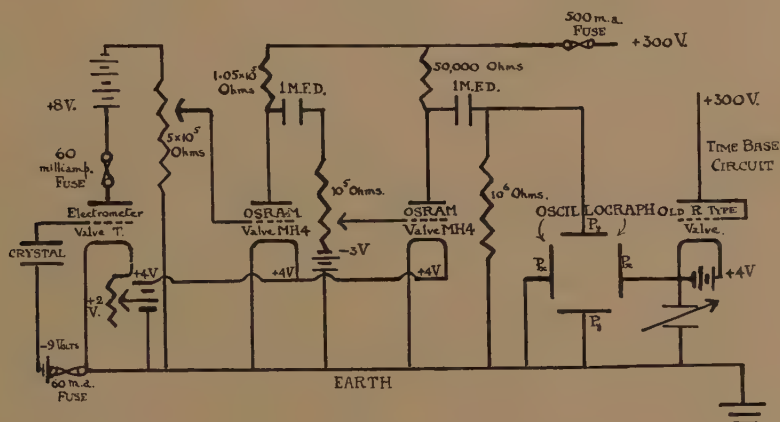
above assumption the area of the projected oscillogram is shown to be proportional to the change of momentum. Other assumptions lead to different relations, but as experiment reveals a linear relation between area and change of momentum, it seems legitimate to deduce that the ordinates of the projected oscillograms are proportional to the instantaneous forces. The change of momentum is calculated as follows. By a well-known property of a circle, if R is the radius of the vertical circle described by the C.G. of the bob, H the vertical distance descended, assuming impact to commence when the C.G. is at its lowest point, and a_1 is the horizontal projection or half-chord of the circular arc of descent, then $H(2R-H) = a_1^2$. Hence $H^2 - 2RH + a_1^2 = 0$ and $H = R - (R^2 - a_1^2)^{\frac{1}{2}}$. When a_1^2 is much less than R^2 the contents of the bracket can be expanded giving $H = a_1^2/2R$, approximately. In our experiments $R = 266.7$ cm., and the largest value of a_1 was 32 cm. The exact formula gave $H = 1.93$ cm., and the approximate one $H = 1.92$ cm. The largest error involved is using the approximate formula is thus about 0.5 per cent. Since non-uniform focusing, origin distortion, and other causes probably produce errors in measuring the areas of the order ± 2 per cent., it has been regarded as permissible to use the approximate formula in the calculations of momenta. After falling from rest under gravity through a height H the velocity at the bottom is $(2gH)^{\frac{1}{2}}$, neglecting air resistance. The momentum just before impact commences is $m(2gH)^{\frac{1}{2}}$, where m is the mass of the bob. This reduces to $m(g/R)^{\frac{1}{2}}a_1$, since $H = a_1^2/2R$, approx., where R and a_1 have the above meanings. Similarly, if the bob finally parts company with the struck string while the string is at its lowest point, and recedes along an arc of a vertical circle of the same radius R , such that the horizontal half-chord is a_2 , the momentum just after contact ceases is $m(g/R)^{\frac{1}{2}}a_2$. Hence the change of momentum is $m(g/R)^{\frac{1}{2}}(a_1 + a_2)$. Since $m(g/R)^{\frac{1}{2}}$ is a constant, the change of momentum is proportional to $(a_1 + a_2)$, and in testing whether the areas under the projected oscillograms are proportional to changes of momenta during impact, it is sufficient to plot $(a_1 + a_2)$ as the variable representing change of momentum. This is what was done in plotting the graphs of fig. 2.

Experimental Arrangements.

The potential differences produced by the force of impact applied to the piezoelectric crystal in the bob were amplified by a circuit which was nearly the same as that used by Grime⁽¹⁾ (see fig. 3). Thus the circuit includes an Osram T electrometer valve as input valve, combined with two-stage amplification. No formal leakage resistance was inserted in parallel with the condenser formed by the plates of the crystal, but there was a natural leak of about 750 megohms. The piezoelectric crystal was a circular cylinder of Rochelle salt, two inches in diameter and a

quarter of an inch thick, sold by Messrs. Rothermel of London. The cylinder consisted of a so-called "bimorph," *i. e.*, two equal halves cemented together in electrical opposition. One electrode was a sheet of tinfoil sandwiched between the two halves, and the other was formed by sheets of tinfoil, one on each end of the combined cylinder, joined together. This system was so arranged that a certain straight line in it was fixed to supports, and the point of application of forces producing piezoelectric effects was at some distance from that line and was free mechanically. Forces acting parallel to the axis of the cylinder, which was also the "*a*" or electric axis of each half, caused a bending of the bimorph and with it a piezoelectric p.d. Rochelle salt, though more brittle, was selected in preference to quartz because its greatest piezoelectric constant is about 8.1×10^{-5} e.s.u. per dyne, whereas that of the

Fig. 3.



The electric circuit operating the oscillograph.

latter is about 6.9×10^{-8} e.s.u. per dyne. Also the technique employed by Messrs. Rothermel gives crystals which will withstand forces of impact up to 4.4×10^6 dynes per sq. cm. without breaking. The leads connecting the crystal to the amplifier (except those parts forming the pendulum wires) and also that connecting the oscillograph to the amplifier, were carefully screened both electrostatically and electromagnetically. The oscillograph was a Cossor model 3232 A. The full diameter of the screen was approximately 13 cm., but only those figures which appeared in the middle flat part of the screen of diameter 10 cm. were used in constructing the graphs. Photographs of oscillograms were taken by an external camera using Ilford Golden Iso-Zenith plates (H.D. 1400).

The developed plates were placed in a lantern, an image was cast on a sheet of squared paper, and a tracing made, using the lower edge

of the image of the oscillogram for the sake of precision. A straight line representing the time axis was drawn by extrapolating the lower edge of the image of the initial undeflected portion of the curve, and the area between this line and that corresponding to the lower edge of the curve was measured. The time base was of the simple type shown diagrammatically on the right-hand side of fig. 3. Each oscillogram was a transient of the single traverse type (see Pl. III.). The linearity of the time base was demonstrated by replacing the crystal-amplifier system by the local A.C. supply of 50 cycles per second and photographing the curve whose ordinates were A.C. potential differences, and whose abscissæ were deflexions produced by the time base circuit. Uniformity in the wave-length of these "ripples" confirmed the linearity of the time base. One such A.C. curve was photographed on each plate for calibration purposes. No special devices were used to keep the oscillograph spot equally well focused at all points of its course or to remove origin distortion, so that small errors were to be expected from want of perfection in these respects, and the calibration ripple displays a kink due to origin distortion. To bring the oscillograph figure into the middle of the screen it was necessary to commence the impact and to start the horizontal motion of the spot controlled by the time base, at the proper moments. Initially the pendulum bob carrying the crystal and forming the hammer was held at rest in a displaced position by an electromagnet. On breaking the circuit of this magnet the bob was released and at the same time a U-shaped piece of copper wire, destined to close the high tension branch of the time base circuit, was released at another place so as to begin to slide down an inclined plane. When the U-shaped copper reached the bottom of the plane its two ends entered two separate pools of mercury and closed the time base circuit. By adjusting the distance of fall down the inclined plane the moment of closing the time base circuit could be altered relatively to the moment of commencement of the impact, until the oscillograph figure appeared in the middle of the screen. This method ensured fair constancy of position of the curves on the screen, in a series of repeated experiments with constant conditions, provided the horizontal speed of the spot was not too great. An electrical method, using a gas-filled relay valve, gave roughly the same degree of repetition.

As already mentioned, the hammer consisted of the bob of a compound pendulum. The method of suspension was by four long, thin, copper wires (see fig. 1). The bob consisted of a rectangular prism of wood with the crystal system attached to the front end, and a rod of iron fixed centrally into the rear end. The crystal system, which was mounted in a metal protecting sheath as supplied by the makers, was of somewhat irregular form, and this made the bob slightly unsymmetrical with respect to a

vertical plane passing through the centre of mass and the struck point of the wire at the moment of impact. It is hoped to remedy this fault in future work.

Various workers distinguish between "soft" and "hard" hammers. To form a soft hammer we used a pendulum bob of the kind just described, with a striking part, viz., the head of a felt-covered piano hammer, mounted in front. This was attached to the metal plate protecting the crystal by Collins' Elastic Glue (see fig. 1). To form a hard hammer the felt-covered hammer head was replaced by a piece of brass of the shape shown. The two types of hammer gave two characteristic types of oscillograms (see Pl. III.). The "struck string" was a length of No. 18 British S.W.G. steel piano wire mounted on a monochord which was firmly clamped with the wire vertical. The direction of motion of the hammer during the experiment was normal to the string and such as to press the string more firmly against its supporting bridges, not to make it slide over them.

Experimental Data.

Diameter of the struck string	= 0.122 cm.
Length of the struck string	= 101.7 cm.
Mass of the string	= 9.27 gm.
Fundamental period of transverse vibrations of the string on 15/8/1938	= 1/182 seconds.
Mass of the soft hammer	= 286.0 gm.
Mass of the hard hammer	= 290.1 gm.
Radius of the circle described by the C.G. of the bob	= 266.7 cm.

Results. Soft Hammer. (See fig. 2.)

In the following tables are given values of (1) momentum changes during impact, represented by $a_1 + a_2$, measured in cm., each cm. representing 548.6 c.g.s. units of momentum, *i. e.*, 548.6 gm. cm. sec.⁻¹; (2) corresponding areas magnified and projected on squared paper by means of a lantern. Each arbitrary unit of area represents 1.61 sq. cm. The graphs of fig. 2 are obtained by plotting these values, areas as abscissæ, momenta as ordinates.

TABLE I.

Set 1 (taken 12/6/1937). Distance of struck point from the near end of the string (the striking length) = 11.8 cm.

$a_1 + a_2$	14.3	17.6	16.5	19.0	31.0	39.0	52.0	63.0	cm.
Area	12.0	15.5	20.0	19.0	27.0	35.0	46.0	56.0	arbitrary units.

Set 2 (taken 24/8/1937). Striking length = 14.8 cm.

$a_1 + a_2$	18.0	22.1	27.2	31.0	40.2	43.4	48.2 cm.
Area	12.5	14.8	19.0	20.5	29.5	31.0	35.5 arbitrary units.

Set 3 (taken 6/11/1937). Striking length = 17.3 cm.

$a_1 + a_2$	17.5	22.4	26.0	30.5	33.6	36.8	39.9	42.4	45.6	49.3 cm.
Area	3.0	3.55	4.2	5.0	5.1	6.15	7.0	6.8	7.0	7.4 a.u.

Discussion of Results.

1. The graphs of fig. 2 obtained by plotting momentum changes of the soft hammer during impact as ordinates, and areas between the projected and magnified oscillograph curves and the time axis as abscissæ are approximately straight lines passing through the origin. Hence the projected oscillograms appear to be nearly true force-time diagrams, *i. e.*, the lengths of the ordinates measured from the time axis, using the lower edges of both curves and time axis for this purpose, are approximately proportional to the instantaneous forces exerted by the struck string on the hammer. This statement is only true if certain precautions are taken. (a) The oscillogram must be formed in the central flat portion of the screen. Curves that were formed at the side or that were too large yielded points off the straight line graphs. This effect seems to be due to change of sensitivity of the electron beam with increasing obliquity as well as to curvature of the screen. (b) The oscillogram must not be too small or the corresponding point lies off the straight line—this is perhaps due to the fault of the oscillograph called origin distortion and to errors in measuring areas caused by the finite but varying width of the line of the oscillogram. An occasional point came right off the straight line, *e. g.*, the point marked A in fig. 2, Set 1, due to incorrect setting of the pendulum while working in a dark room.

2. The general shape of the oscillograms or force-time diagrams is that of a single hump, with numerous small indentations superposed on it. The force starts from value zero at the beginning of impact, rises to a maximum, and falls away to zero again when contact ceases. The process takes a finite time. Using the terminology of acoustics we may say that the force exerted by the string on the hammer consists of a fundamental term accompanied by a suite of overtones. With the hard hammer the overtones are much more intense relative to the fundamental, much sharper and more complex than with the soft hammer (see Pl. III.). The main hump of the curves is probably due to the primary direct force exerted by the string upon the hammer during impact. The problem of the full analysis of the overtones is reserved for a later paper, but among the probable causes of them may be mentioned (a) transverse waves on the string which, having started from the struck point, have travelled in each direction and been reflected at the ends. Such waves

repeatedly return to the struck point, where they exert a certain force on the hammer so long as contact between hammer and string is maintained. Also every time they reach the struck point they become more complex, since they are partly reflected and partly transmitted. Various workers, especially C. V. Raman⁽²⁾, P. Das⁽³⁾, R. N. Ghosh⁽⁴⁾, K. C. Kar⁽⁵⁾, M. Ghosh⁽⁶⁾, and S. C. Dhar⁽⁷⁾, given theoretical treatments of the struck string dealing with these waves, (b) longitudinal waves in the hammer itself, propagated towards the rear through crystal, wood, and steel, which after reflexion by the final or intermediate surfaces of separation of two media, return to the crystal again, exerting a force on it. This process is likely to be repeated. On account of the high value of the velocity of longitudinal waves in crystal, wood, steel, etc., waves describing the full return length of the hammer would do so in about 10^{-1} sec., causing indentations with peaks about 10^{-4} mm. apart on fig. D of Pl. III. Hence the more striking indentations of figs. E and F of Pl. III. are not due to these longitudinal waves as they are too far apart. (c) Electrical oscillations of the crystal regarded as a Butterworth-Dye equivalent circuit are too rapid to cause the indentations. Electrical oscillations in a circuit formed by the crystal regarded as a condenser in series with its leakage resistance and a self-inductance would cause corresponding indentations on the oscillograms. In the present case, however, the capacity is of the order 8×10^{-10} farads, the leakage resistance 7.5×10^8 ohms, and the self-inductance greater than 1.7×10^{-5} henrys, and the circuit is non-oscillatory. (d) Sound waves in the atmosphere coming from the sound board of the monochord, the bridges and the other objects may contribute to the complexity of the oscillograms by exerting periodic forces on the hammer head. These would be expected to have a period equal to that of the fundamental note of the string and its overtones.

Of the four types of wave suggested, type (a) would appear to be most likely to cause the indentations. Examples of the relatively less complex oscillograms obtained with the soft hammer are seen in Pl. III. figs. A, B, C, and D. Other similar oscillograms have been used in obtaining the straight line graphs of fig. 2. Fig. D of Pl. III. has been subjected to harmonic analysis. A magnified image of the lower edge of the curve of fig. D (Pl. III.) was analysed by means of a Mader-Ott harmonic analyser. It was found that y , the ordinate of any point on the lower edge of the curve of fig. D (Pl. III.) could be expressed in terms of x , its abscissa, measured from the point representing the beginning of impact, by the series :—

$$y = 1.28 \{ 1.00 \sin \pi x/b - 0.04 \sin 2\pi x/b - 0.07 \sin 3\pi x/b \\ + 0.00 \sin 4\pi x/b - 0.01 \sin 5\pi x/b \} \text{ in cm.}$$

The sixth and other components up to the ninth, which was the highest

the analyser could reveal, had amplitudes less than 1 per cent. of the first. In the series b represents the length of the base of the main curve, *i. e.*, the length of the intercept which it makes with the time axis. The first component is seen to be by far the most important. Visual inspection suggests that the small indentations in fig. D (Pl. III.) could be regarded as due to, say, a fifteenth to an eighteenth component in the above series. The harmonic analysis of hard hammer oscillograms has not been attempted, as the indentations judged visually appear to correspond to components of a higher order than the ninth. Hence no algebraic expression has been obtained in this case.

3. In the original of fig. D (Pl. III.) a length of 1.71 cm. corresponds to a time interval of $1/50$ sec. Now the base of the figure representing the duration of impact is 2.75 cm. long, approx., which therefore represents 3.2×10^{-2} secs. Also since $x = k_1 t$ and when $x = 1.71$, $t = 1/50$ sec., it follows that $k_1 = 85.5$ cm. per sec. The area of the original of fig. D (Pl. III.), between the lower edges of the curve and time axis, as measured by a planimeter, is 2.30 sq. cm. Now $y = k_2 F$ and the area is $k_1 k_2 \int F dt = k_1 k_2 \times \text{momentum change}$. The momentum change is found as above to be 1.05×10^4 gm. cm. sec.⁻¹. Hence $k_1 k_2 = 2.30 / (1.05 \times 10^4)$, and since $k_1 = 85.5$, k_2 is found to be 2.57×10^{-6} cm. per dyne, $1/k_2$ is 3.89×10^5 dynes per cm. Thus in the original curve of fig. D (Pl. III.) an ordinate of 1 cm. represents 3.89×10^5 dynes. The equation to the lower edge of the oscillogram is $y = 1.28 \sin \pi x/b - 0.05 \sin 2\pi x/b - 0.09 \sin 3\pi x/b$ approx., where $b = 2.75$. This gives an instantaneous force

$$F = y/k_2 = 5.0 \times 10^5 \sin \pi x/b - 1.9 \times 10^4 \sin 2\pi x/b \\ - 3.5 \times 10^4 \sin 3\pi x/b \text{ dynes.}$$

The measured maximum y is 1.40 cm., agreeing well with the value 1.37 given by the formula. The measured maximum force is about 5.45×10^5 dynes.

4. In the above experiments the ratio mass of hammer/mass of string was roughly as thirty-one to one. In all or nearly all existing theories of the struck string a much smaller mass ratio is assumed, so that the present work cannot be used to test those theories.

5. At the end of each impact the oscillograph spot descended a short distance below the time axis, to which it gradually returned, all the while describing small oscillations. This effect has not been fully investigated, but it seems reasonable to suppose that when a piezoelectric crystal, initially under compression, is released so that it relaxes rapidly, electric charges of opposite sign to those produced by compression will appear and manifest themselves by some such effect as this.

6. It is hoped to remove the sources of error called origin distortion and non-uniform focusing of the oscillograph spot by using a more

recent type of instrument. A new, symmetrical, but light pendulum bob has been designed. A much longer string is being mounted, so that theories of the struck string can be tested.

We desire to express our hearty thanks to Professor L. F. Bates for his continued interest and support, to Dr. W. H. George for a very profitable discussion, to Mr. H. Watson for help in constructing apparatus, and to Mr. C. Calam for the loan of a photographic lens.

References.

- (1) G. Grime, *Proc. Phys. Soc.* xlv. p. 196 (1936).
- (2) C. V. Raman and B. Banerji, *Proc. Roy. Soc. A*, xcvi. p. 99 (1920).
- (3) P. Das, *Proc. Indian Assoc. Cult. Sci.* vii. p. 13 (1921-22); ix. p. 297 (1925-26); x. pp. 75 and 437 (1926-27). N.B.—Vol. x. of *Proc. Indian Assoc. Cult. Sci.* is also Vol. i. of *Indian Journ. Phys.*). P. Das, *Proc. Phys. Soc.* xl. p. 31 (1928). P. Das and S. K. Datta, *Phil. Mag.* vi. p. 479 (1928).
- (4) R. N. Ghosh and S. Bhargava, *Phil. Mag.* xlvii. p. 1141 (1924), and xlix. p. 121 (1925). R. N. Ghosh, *Phil. Mag.* ix. p. 1174 (1930); *Phys. Rev.* xxiv. p. 456 (1924); *Proc. Phys. Soc.* xli. p. 224 (1928); *J. Acoust. Soc. America*, vii. p. 254 (1936). R. N. Ghosh and H. G. Mohammad, *Phil. Mag.* xix. p. 305 (1935).
- (5) K. C. Kar, *Phil. Mag.* vi. p. 276 (1928); *Phys. Rev.* xxi. p. 695 (1923). K. C. Kar and M. Ghosh, *Phil. Mag.* ix. pp. 306 and 321 (1930) and xii. p. 676 (1931).
- (6) M. Ghosh, *Indian Phys.-Math. J.* iii. p. 15 (1932); *Phil. Mag.* xvii. p. 521. (1934).
- (7) S. C. Dhar, *Indian Journ. Phys.* x. p. 305 (1936).

*XIV. The Variation of Solar Ultra-Violet Radiation
during the Sunspot Cycle *.*

By E. V. APPLETON, F.R.S., and R. NAISMITH, A.M.I.E.E. †.

[Received January 7, 1939.]

(1) *Introductory.*

THE first measurements of upper-atmospheric ionization employing the critical penetration frequency method were concerned with Region E of the ionosphere ‡. Such measurements were at first made at somewhat irregular intervals during 1930, while the technique of the method was being elaborated, but from the beginning of 1931, there is now available a systematic series of observations of eight years' duration. To such measurements of the ionization density in Region E there can now be added the additional data for Regions F_1 and F_2 , obtained both by ourselves, and also by other workers, in different parts of the world, who have adopted the critical frequency method.

Since there is now ample evidence for the conclusion that the ionosphere is caused by solar radiation, it is of interest to examine the variation of ionization over a period which is now sufficiently extensive to include years of both sunspot maximum and sunspot minimum in order to estimate how the intensity of the relevant solar radiation has itself varied during the interval. In this communication the results of such an examination are described, consideration being restricted to the cases of Regions E and F_1 .

(2) *The Relation between Ionizing Radiation Intensity
and Ionospheric Electron Density.*

We have previously shown §, from a series of noon measurements of the ionization in Regions E and F_1 , that the electrons in these regions disappear according to a recombination law. This has more recently

* Paper presented in abstract at the Sixth General Assembly of the International Scientific Radio Union, Venice, Sept. 1938.

† Communicated by Professor E. V. Appleton, F.R.S.

‡ Appleton, "A Method of Measuring Upper-Atmospheric Ionization," 'Nature' (Feb. 7th, 1931). Appleton and Naismith, Proc. Roy. Soc. A, cxxxvii. p. 36 (1932).

§ Appleton and Naismith, Proc. Roy. Soc. A, cl. p. 685 (1935).

been confirmed by Kirby, Gilliland, and Judson * in observations made at Washington during the solar eclipse of February 3rd, 1935. We can write, therefore, for stationary conditions, such as exist at noon,

$$q = \alpha N^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where q is the effective rate of electron production, α is the effective electron recombination coefficient, and N the electron density. Now the relation between the maximum electron density N and the ordinary wave critical penetration frequency f_0 for any region may be written

$$N = K f_0^2 \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot (2)$$

where K is a known numerical constant. Moreover, the maximum rate of electron production in a region (g) is related to the sun's zenith distance (χ) by the relation

$$q = q_0 \cos \chi, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where q_0 is the maximum rate of ion production at the equator at the equinoxes.

From equations (1), (2), and (3) we have

$$\frac{f_0^4}{\cos \chi} = \frac{q_0}{\alpha K^2} \dots \dots \dots (4)$$

It is proposed that the quantity on the left-hand side of (4) should be taken as the *Region character figure*, since, if α is independent of the air-pressure at the level at which the ionized region is formed, such a figure should have the same value at all stations and be proportional to the intensity of the solar radiation responsible for the electron production in the particular region with which we are concerned. For convenience f_0 may be expressed numerically in megacycles per second.

(A word should here be interpolated concerning the physical significance of the effective recombination coefficient α and also of the quantity q_0 . Following Massey [†], who has shown that the processes of attachment of electrons to and detachment from neutral atoms is such as to form a quasi-stationary balance between the negative ion and electron concentrations, Appleton and Sayers [‡] have, in a recent paper, shown that the effective recombination coefficient for electrons is equal to $(\alpha_e + \lambda\alpha_i)$, where α_e is the recombination coefficient for electrons and positive ions, α_i is the recombination coefficient for negative ions and positive ions, and λ is the ratio of the negative ion and electron concentrations. They have

* Kirby, Gilliland and Judson, P.I.R.E. xxiv. p. 1027 (1936).

† Massey, Proc. Roy. Soc. A, clxiii. p. 542 (1937).

† Appleton and Sayers, Sixth General Assembly, U.R.S.I. 'Proceedings,' p. 272 (1938).

similarly shown that the effective rate of electron production per c.c. q_0 is equal to $q_0'/1+\lambda$, where q_0' is the actual rate.)

(3) *Variation of Region Character Figure with Sunspot Cycle.*

As a result of the two years' measurements of noon ionization for Region E in 1931 and 1932, we noticed that the general ionization level was less in the latter year than in the former. We pointed out that it was natural to associate this with the general fall in solar activity in the sequence of the solar cycle and further added the following sentence: "If work in future years confirms this variation with the sunspot cycle and we are able to ascribe it even only in part to the ultra-violet light from the Sun, it will follow that such radiation varies with solar activity over a wider range than does the so-called solar constant." As will be shown below, we are now able to confirm our earlier view concerning the sympathetic variation of ionization with solar activity, and since it has, in the meantime, been shown that the chief ionizing agency for Region E is

TABLE I.

	1931.	1932.	1933.	1934.	1935.	1936.	1937.	1938.
Region E	237·3	201·9	213·2	134·9	186·7	263·3	337·4	291·3
Region F ₁	—	—	426·7	474·7	487·6	898·0	1080	898·0

solar ultra-violet light, the consequence stated in our earlier paper necessarily follows.

For the cases of both regions E and F₁, all the available region character figures for the years 1931 to 1938 have been calculated. The summer values for the Slough station in South-East England (Latitude $51\frac{1}{2}^\circ$ N.) are set out in Table I.

Since, however, the region character figure is so defined as to permit its determination at any station, we have also made out a similar table (Table II.) in which all the available data are included by taking the mean of the summer character figures obtained from observations at Tromsø (Norway) (Lat. $69\frac{1}{2}^\circ$ N.), Slough (Lat. $51\frac{1}{2}^\circ$ N.) (England), and Washington (Lat. $48\frac{1}{2}^\circ$ N.) (U.S.A.).

It will be seen that the general mean results do not differ markedly from those of the Slough series.

In fig. 1 the Slough region character figures for summer are plotted, and a smooth curve drawn through them. It will be seen that the minimum electron production occurred in 1933/34, while the maximum was reached in 1937/38. Now the sunspot minimum for this particular

period is known to have occurred at 1933·8, while the sunspot maximum of 1937–38 has just passed.

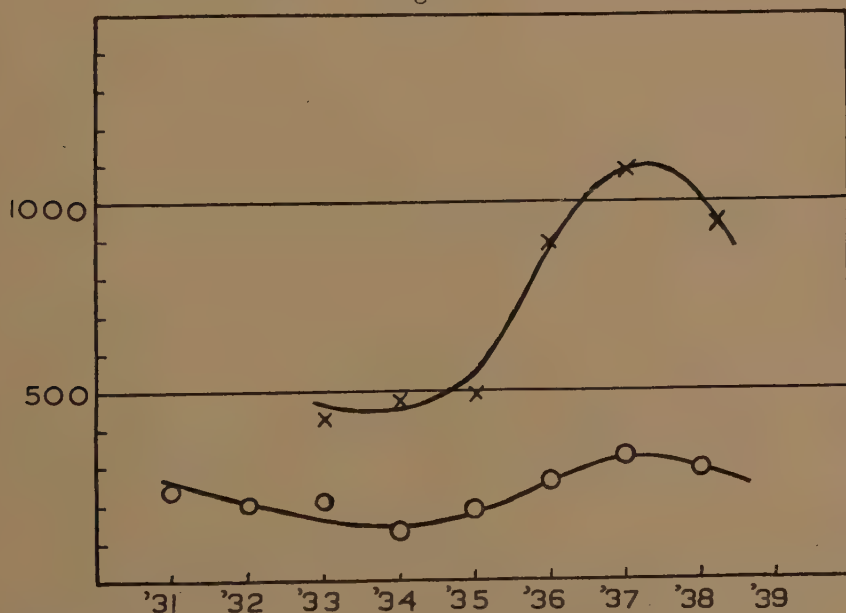
From fig. 1 it is possible to estimate the ratio of the sunspot maximum and sunspot minimum values of ion-production for the two Regions E and F_1 . These are respectively 2·25 : 1 and 2·45 : 1. We thus conclude that

TABLE II.

	1931.	1932.	1933.	1934.	1935.	1936.	1937.	1938.
Region E	237·3 (S.)	201·9 (S.)	213·2 (S.)	129 (S.W.)	182	249	297	291·3 (S.)
Region F_1	—	—	426·7 (S.)	414·8 (S.W.)	449	841·3	1106	967·7 (S.)

S.=Slough alone. S.W.=Slough and Washington only.

Fig. 1.



Region character figures for Regions E and F_1 during period 1931–38.
(O Region E; X Region F_1 .)

the absorbed ultra-violet radiation causing Regions E and F_1 has varied enormously during the sunspot cycle compared with any variation there may be in the intensity of the radiation received at ground-level.

It will be noticed that the ionization in Region E has varied by 50 per cent. and in Region F_1 by 56 per cent. during the cycle, so that the direct-current conductivity of these regions has increased correspondingly. It may be mentioned in this connexion that it has been known for many years that the solar magnetic diurnal variation is about 60 per cent. greater at sunspot maximum than at sunspot minimum.

SUMMARY.

Measurements of upper-atmospheric ionization by the critical frequency method show that the solar ionizing agent (ultra-violet light) responsible for the formation of Regions E and F_1 of the ionosphere varies by 120 to 150 per cent. during the sunspot cycle. The ionization and electrical conductivities of these regions vary correspondingly by 50 to 60 per cent.

XV. *Euler's Problem of Two Fixed Centres of Gravitation.*

By MONROE H. MARTIN †.

[Received September 1, 1938.]

INTRODUCTION.

THE purpose of this paper is to classify the various forms of motion possible for a material point P moving in a plane under the gravitational attraction of two masses K and K' fixed in this plane. Such a classification has been made by Charlier ‡, whose treatment, however, leaves doubt as to whether all the forms of motion classified by him actually exist. As a matter of fact, it turns out § that his classification includes forms of motion which do not exist, and fails to include another which does exist.

The case where $K=K'$ is considered, and, as might be expected, is somewhat simpler than the case $K \neq K'$. Various forms of motion present when $K \neq K'$ no longer appear when $K=K'$. This fact leads one to suspect that the same state of affairs might be true also in the restricted problem of three bodies in which the masses K and K' are not fixed but revolve in concentric circles. Numerous forms of motion have been established by numerical computation || in the restricted problem of three bodies when the two revolving masses are equal. The behaviour of Euler's problem accordingly leads one to suspect that the restricted problem of three bodies possesses forms of motion not accessible to an attack of this kind.

PART I.—THE GENERAL PROBLEM.

1. *The Differential Equations of Motion.*—The Lagrangian function of Euler's problem in rectangular coordinates (x, y) is given by

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + U,$$

† Communicated by the Author.

‡ C. V. L. Charlier, 'Die Mechanik des Himmels,' vol. i. (1927). For a bibliography of Euler's problem see 'Encyklopadie der Mathematischen Wissenschaften,' Bd. 4. pt. 2 p. 497.

§ See § 6, p. 171.

|| E. Strömgren has published the results of an enormous amount of work along these lines. See, for example, his paper "Forms of Periodic Motion in the Restricted Problem and in the General Problem of Three Bodies according to Researches executed at the Observatory of Copenhagen," no. 39 of 'Publikationer og mindre Meddelelser fra Københavns Observatorium.'

where

$$U = K/r + K'/r', \quad r^2 = (x-c)^2 + y^2, \quad r'^2 = (x+c)^2 + y^2,$$

and the masses K' , K ($0 < K' \leq K$) are situated at the points $(-c, 0)$, $(c, 0)$ respectively. These points are conveniently designated by K' and K . The Lagrangian equations of motion are

$$\ddot{x} = U_x, \quad \ddot{y} = U_y, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and possess the energy integral

$$\frac{1}{2}(\dot{x}^2 + \dot{y}^2) - U = h, \quad (h = \text{const.}) \quad . \quad . \quad . \quad . \quad . \quad (2)$$

From (1) it appears that the dynamical system is reversible, and from (2) it is obvious that any motion must occur within the curve of zero velocity $-U = h$. An easy analysis of the curves of zero velocity for various values of h shows that for $h < 0$ the motion is restricted to a finite part of the plane, while for $h \geq 0$ this restriction no longer holds.

On introducing confocal coordinates (u, v) by setting

$$r + r' = 2u, \quad r - r' = 2v, \quad (u \geq c, \quad |v| \leq c),$$

the Lagrangian function takes the well-known form of Liouville,

$$L = \frac{1}{2}\phi(A\dot{u}^2 + B\dot{v}^2) + U,$$

where

$$U = f/\phi, \quad f = a(u) + b(v), \quad \phi = \alpha(u) + \beta(v),$$

$$A = A(u), \quad B = B(v).$$

In the problem at hand we have

$$\left. \begin{aligned} A(u) &= \frac{1}{u^2 - c^2} > 0, \quad a = a(u) = (K + K')u, \quad \alpha = \alpha(u) = u^2, \\ B(v) &= \frac{1}{c^2 - v^2} > 0, \quad b = b(v) = -(K - K')v, \quad \beta = \beta(v) = -v^2 \end{aligned} \right\} \quad . \quad . \quad (3)$$

From (1) it may be seen that motion for which $\dot{y} \equiv 0$ is possible. For such motion $u \equiv c$ or $v \equiv \pm c$, and this motion we rule out for the time being, since A and B above are not defined. We shall return to this rectilinear motion in § 5.

The Lagrangian equations of motion are

$$\left. \begin{aligned} \frac{d}{dt}(\phi A \dot{u}) - \frac{1}{2}\phi_u(A\dot{u}^2 + B\dot{v}^2) - \frac{1}{2}\phi \frac{dA}{du}\dot{u}^2 &= U_u, \\ \frac{d}{dt}(\phi B \dot{v}) - \frac{1}{2}\phi_v(A\dot{u}^2 + B\dot{v}^2) - \frac{1}{2}\phi \frac{dB}{dv}\dot{v}^2 &= U_v, \end{aligned} \right\} \quad . \quad . \quad . \quad . \quad (4)$$

and possess the energy integral

$$\frac{1}{2}\phi(A\dot{u}^2 + B\dot{v}^2) - U = h. \quad . \quad . \quad . \quad . \quad (5)$$

Multiplying both equations in (4) by ϕ , and using (5), they may be written

$$\left. \begin{aligned} \phi \frac{d}{dt} (\phi A \dot{u}) - \frac{1}{2} \phi^2 \frac{dA}{du} \dot{u}^2 &= \frac{d}{du} (a + h\alpha), \\ \phi \frac{d}{dt} (\phi B \dot{v}) - \frac{1}{2} \phi^2 \frac{dB}{dv} \dot{v}^2 &= \frac{d}{dv} (b + h\beta). \end{aligned} \right\} \dots \dots \dots (6)$$

Multiplying the first equation in (6) by \dot{u} , the second by \dot{v} , and then integrating †, we arrive at the well-known equations

$$\phi^2 A \dot{u}^2 = 2(h\alpha + a + k), \quad \phi^2 B \dot{v}^2 = 2(h\beta + b - k), \quad (k = \text{const.}). \quad (7)$$

Let

$$\left. \begin{aligned} L(u) &= h\alpha + a + k, & R(u) &= 2(u^2 - c^2)L(u), \\ M(v) &= -h\beta - b + k, & S(v) &= 2(v^2 - c^2)M(v), \end{aligned} \right\} \dots \dots \dots (8)$$

and equations (7) may be written, after substituting for A, B, α, β their values given in (3), in the form

$$(u^2 - v^2)^2 \dot{u}^2 = R(u), \quad (u^2 - v^2)^2 \dot{v}^2 = S(v),$$

which may also be written

$$\dot{u} = \frac{\sqrt{R(u)}}{u^2 - v^2}, \quad \dot{v} = \frac{\sqrt{S(v)}}{u^2 - v^2}, \quad \dots \dots \dots (9)$$

wherein $\sqrt{R(u)}, \sqrt{S(v)}$ take the sign of \dot{u}, \dot{v} respectively.

Further integration yields

$$0 = \int_{u_0}^u \frac{du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{dv}{\sqrt{S(v)}}, \quad t - t_0 = \int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{v^2 dv}{\sqrt{S(v)}}, \quad (10)$$

so that $t = t_0$ when $u = u_0, v = v_0$. In (10), as in (9), the radicals $\sqrt{R(u)}, \sqrt{S(v)}$ take the signs of \dot{u}, \dot{v} respectively. It follows that each integral in (10) always has a positive value. The differential equations of motion are now integrated, and we take up the study of the qualitative nature of the motion.

2. *Motion in the (u, v) -plane.*—The motion in the (u, v) -plane is represented by the movement of a point (u, v) on a curve

$$C: \quad u = u(t), \quad v = v(t),$$

drawn in the region

$$u \geq c, \quad |v| \leq c.$$

It appears from (9) that motion is possible only in a region

$$R: \quad R(u) \geq 0, \quad S(v) \geq 0,$$

† This operation introduces extraneous solutions which we guard against later in Lemma 4 of § 2.

and we shall investigate motion in these regions for $t > t_0$. Since the dynamical system is reversible, the forms of motion for $t < t_0$ are identical with those for $t > t_0$. The regions R are rectangles (one side of which may lie at infinity) whose sides are parallel to the coordinate axes in the (u, v) -plane. In special instances the rectangle R may degenerate into a straight line segment of one of the following kinds :—

$$\begin{aligned} u &= u_0, & |v| &\leq c, \\ u &\geq c, & v &= v_0, \\ c &\leq u \leq u_0, & v &= v_0, \end{aligned}$$

where $u_0 > c$, $|v_0| < c$.

We consider first the case of a non-degenerate rectangle R . In the interior of R both $R(u)$ and $S(v)$ are positive. On the boundaries of R exactly one of $R(u)$, $S(v)$ vanishes except at the corners of the rectangle, where both vanish. The right-hand members of (9) are analytic throughout the interior of R , but are not analytic on its boundary. The curve C consists of a number of analytic arcs joined together at their endpoints on the boundary of R . On each of these arcs $u(t)$, $v(t)$ are monotone functions of t , and we shall now prove a series of lemmas to determine the nature of these functions at the boundary points.

LEMMA 1.—*The system of differential equations*

$$\dot{u} = \sqrt{u - u_0} P(u, v), \quad \dot{v} = Q(u, v),$$

where $P(u, v)$, $Q(u, v)$ denote real power series in the arguments $u - u_0$, $v - v_0$ convergent in a given neighbourhood of (u_0, v_0) with $P(u_0, v_0) \neq 0$, $Q(u_0, v_0) \neq 0$, possesses a real power series solution passing through (u_0, v_0) of the form

$$u - u_0 = \alpha^2 t^2 + \dots, \quad v - v_0 = \beta t + \dots, \quad (\alpha \neq 0, \beta \neq 0),$$

valid for $|t|$ sufficiently small. Excluding the case $u \equiv u_0$, the solution through (u_0, v_0) is unique.

Introducing new variables \bar{u} , \bar{v} by means of the equations

$$u = u_0 + \bar{u}^2, \quad v = v_0 + \bar{v},$$

the differential equations become, after dividing out the factor \bar{u} ,

$$\dot{\bar{u}} = \bar{P}(\bar{u}, \bar{v}), \quad \dot{\bar{v}} = \bar{Q}(\bar{u}, \bar{v}),$$

where $\bar{P}(\bar{u}, \bar{v})$, $\bar{Q}(\bar{u}, \bar{v})$ are real power series convergent for sufficiently small $|\bar{u}|$, $|\bar{v}|$ and $\bar{P}(0, 0) \neq 0$, $\bar{Q}(0, 0) \neq 0$. It follows from elementary existence theorems that this system possesses a unique real power series solution of the form

$$\bar{u} = \alpha t + \dots, \quad \bar{v} = \beta t + \dots, \quad (\alpha \neq 0, \beta \neq 0),$$

and the stated result follows on returning to the original variables u, v .

LEMMA 2.—The system of differential equations

$$\dot{u} = \sqrt{u - u_0} P(u, v), \quad \dot{v} = \sqrt{v - v_0} Q(u, v),$$

where $P(u, v)$ denote power series like those in Lemma 1, possesses a real power series solution passing through (u_0, v_0) of the form

$$u - u_0 = \alpha^2 t^2 + \dots, \quad v - v_0 = \beta^2 t^2 + \dots, \quad (\alpha \neq 0, \beta \neq 0),$$

valid for $|t|$ sufficiently small. Excluding the cases in which $u \equiv u_0$ or $v \equiv v_0$, the solution through (u_0, v_0) is unique.

After introducing new variables by means of the equation

$$u = u_0 + \bar{u}^2, \quad v = v_0 + \bar{v}^2,$$

the proof of this lemma is the same as the proof of Lemma 1.

The above lemmas hold with obvious modifications if the radicals $\sqrt{u - u_0}$, $\sqrt{v - v_0}$ are replaced by the radicals $\sqrt{u_0 - u}$, $\sqrt{v_0 - v}$ respectively.

LEMMA 3.—The system of differential equations

$$\dot{u} = \frac{\sqrt{u - c}}{u + v} P(u, v), \quad \dot{v} = \frac{\sqrt{v + c}}{u + v} Q(u, v), \quad P(c, -c) \neq 0, \quad Q(c, -c) \neq 0,$$

where $P(u, v)$, $Q(u, v)$ denote real power series in $u - c$, $v + c$ convergent in a given neighbourhood of $(c, -c)$, possesses a real solution passing through $(c, -c)$ of the form

$$u - c = R^2(t^{1/3}), \quad v + c = S^2(t^{1/3}), \quad R(0) = 0, \quad S(0) = 0,$$

where $R(t^{1/3})$, $S(t^{1/3})$ denote power series in $t^{1/3}$. Excluding the cases where $u \equiv c$ or $v \equiv -c$, the solution through $(c, -c)$ is unique †.

Let
$$u - c = \bar{u}^2, \quad v + c = \bar{v}^2,$$

and the differential equations take the form

$$\bar{u} = \frac{\bar{P}(\bar{u}, \bar{v})}{\bar{u}^2 + \bar{v}^2}, \quad \dot{\bar{v}} = \frac{\bar{Q}(\bar{u}, \bar{v})}{\bar{u}^2 + \bar{v}^2}, \quad \bar{P}(0, 0) \neq 0, \quad \bar{Q}(0, 0) \neq 0,$$

where $\bar{P}(\bar{u}, \bar{v})$, $\bar{Q}(\bar{u}, \bar{v})$ denote power series in \bar{u} , \bar{v} convergent for sufficiently small $|\bar{u}|$, $|\bar{v}|$. Dividing these equations yields

$$\frac{d\bar{v}}{d\bar{u}} = \frac{\bar{Q}(\bar{u}, \bar{v})}{\bar{P}(\bar{u}, \bar{v})} = \bar{R}(\bar{u}, \bar{v}),$$

† This lemma is applied in the text to the "orbits of ejection." For a discussion of such orbits see Strömgren, *loc. cit.*, or E. Picard, 'Traité D'Analyse,' t. 3 (1928), pp. 601–608, where further references are given. See also G. D. Birkhoff, "The Restricted Problem of Three Bodies," *Rendiconti del Circolo Matematico di Palermo*, t. 39 (1915), and T. Levi-Civita, "Traiettorie singolari ed urti nel problema ristretto dei tre corpi," *Annali di Matematico*, sér. iii. t. 9 (1903).

where $\bar{R}(\bar{u}, \bar{v})$ is a power series in \bar{u}, \bar{v} of the same type as $\bar{P}(\bar{u}, \bar{v}), \bar{Q}(\bar{u}, \bar{v})$. This differential equation possesses a single power series solution

$$\bar{v} = \alpha \bar{u} + \dots, \quad (\alpha \neq 0), \quad \dots \quad (11)$$

passing through the origin $\bar{u} = \bar{v} = 0$. In order to express \bar{u}, \bar{v} in terms of the time t substitute the above power series for \bar{v} in the first equation of the preceding system. One obtains a differential equation of the type

$$\dot{\bar{u}} = \bar{S}(\bar{u})/\bar{u}^2,$$

where $\bar{S}(\bar{u})$ denotes a power series in \bar{u} with $S(0) \neq 0$. Integration yields

$$t = \beta \bar{u}^3 + \dots, \quad (\beta \neq 0),$$

the integration constant being chosen so that $u = 0$ when $t = 0$. Hence

$$\bar{u} = R(t^{1/3}), \quad (R(0) = 0),$$

where $R(t^{1/3})$ denotes a power series in $t^{1/3}$. Likewise

$$\bar{v} = S(t^{1/3}), \quad (S(0) = 0),$$

and the statement in the lemma follows upon returning to the original variables u, v .

Replacing v by $-v$ in the above lemma, we obtain the following corollary:—

COROLLARY.—*The system of differential equations*

$$\dot{u} = \frac{\sqrt{u-c}}{u-v} P(u, v), \quad \dot{v} = \frac{\sqrt{c-v}}{u-v} Q(u, v), \quad P(c, c) \neq 0, \quad Q(c, c) \neq 0,$$

where $P(u, v), Q(u, v)$ denote power series in $u-c, c-v$ convergent in a given neighbourhood of (c, c) possesses a real solution passing through (c, c) of the form

$$u-c = R^2(t^{1/3}), \quad c-v = S^2(t^{1/3}), \quad R(0) = 0, \quad S(0) = 0,$$

where $R(t^{1/3}), S(t^{1/3})$ denote power series in t . Excluding the cases where $u \equiv c$ or $v \equiv c$, the solution through (c, c) is unique.

Let $r_1, r_2 (c \leq r_2 < r_1)$ denote two consecutive roots of $R(u) = 0$ between which $R(u) > 0$ and let $s_1, s_2 (-c \leq s_2 < s_1 \leq c)$ be two consecutive roots of $S(v) = 0$ between which $S(v) > 0$. Consider a motion lying in the rectangle

$$R: \quad r_2 \leq u \leq r_1, \quad s_2 \leq v \leq s_1,$$

and passing through the interior point (u_0, v_0) of R when $t = t_0$. Such motions occur when $h < 0$.

It will be sufficient to distinguish four cases.

Case I.— r_1, r_2, s_1, s_2 simple roots. To fix the ideas suppose $\dot{u}_0 < 0, \dot{v}_0 < 0$. The motion moves monotonely to the left and downwards, tending towards a point on the boundary of R as t increases. Since the roots of

$R(u)=0$, $S(v)=0$ are all simple, it follows from the second equation in (10) that the motion reaches the boundary after a finite lapse of time. Suppose that it meets $u=r_2$ before meeting $v=s_2$. It appears from Lemma 1 that the motion touches $u=r_2$ and then proceeds monotonely to the right. In case the motion meets $v=s_2$ before meeting $u=r_2$ only obvious modifications are necessary. When the motion meets the vertex (r_2, s_2) of R and this vertex is not the point $(c, -c)$ we may apply Lemma 2, inferring from this lemma that after meeting the vertex (r_2, s_2) the motion returns to the interior of R . In case the motion meets the vertex (r_2, s_2) of R and this vertex is the point $(c, -c)$ the motion in the (u, v) -plane corresponds to a collision in the (x, y) -plane of the material point P with the mass K . We infer from Lemma 3 that after meeting the vertex (r_2, s_2) of R the motion in the (u, v) -plane returns to the interior of R .

Accordingly, no matter which of the four above possibilities occurs, the motion returns to the interior of R after a finite lapse of time, and the above argument, with appropriate modifications due to changes in initial conditions, may be repeated indefinitely. It is readily seen on the basis of the preceding three lemmas and the corollary that the curve C upon which motion takes place contains only ordinary points, save when it meets a vertex of R where it may possess a singular point.

Let

$$\begin{aligned}\mu_1 &= \int_{r_2}^{r_1} \frac{du}{\sqrt{R(u)}} > 0, & \nu_1 &= \int_{s_2}^{s_1} \frac{dv}{\sqrt{S(v)}} > 0, \\ \mu_2 &= \int_{r_1}^{r_2} \frac{u^2 du}{\sqrt{R(u)}} > 0, & \nu_2 &= \int_{s_1}^{s_2} \frac{v^2 dv}{\sqrt{S(v)}} > 0.\end{aligned}$$

These quantities are all finite, and equations (10) may be written in the form

$$\begin{aligned}0 &= \int_{u_0}^u \frac{du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{dv}{\sqrt{S(v)}} + 2m\mu_1 - 2n\nu_1, \\ t - t_0 &= \int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{v^2 dv}{\sqrt{S(v)}} + 2m\mu_2 - 2n\nu_2, \quad \dots \quad (10')\end{aligned}$$

where m, n denote positive integers or zero and

$$\int_{u_0}^u \frac{du}{\sqrt{R(u)}} < 2\mu_1, \quad \int_{v_0}^v \frac{dv}{\sqrt{S(v)}} < 2\nu_1, \quad \int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} < 2\mu_2, \quad \int_{v_0}^v \frac{v^2 dv}{\sqrt{S(v)}} < 2\nu_2.$$

In order that $t - t_0$ tend to $+\infty$ it is clearly necessary that the integer m in the second equation of (10') tend to $+\infty$; from the first equation in (10') it follows that n likewise tends to $+\infty$. Consequently it is not at all obvious from the second equation in (10') that $t - t_0$ tends to $+\infty$

Case III.— r_1, r_2, s_1 simple roots, s_2 a non-zero multiple root. Here μ_1, μ_2 are both finite and ν_1, ν_2 both infinite. Equations (10') are replaced by

$$0 = \int_{u_0}^u \frac{du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{dv}{\sqrt{S(v)}} + 2m\mu_1,$$

$$t - t_0 = \int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{v^2 dv}{\sqrt{S(v)}} + 2m\mu_2. \quad (10''')$$

In order that $t - t_0$ tend to $+\infty$ the integer m in (10''') must tend to $+\infty$. Therefore from the first of these equations v must tend towards s_2 as m becomes infinite. Now

$$\int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} - \int_{v_0}^v \frac{v^2 dv}{\sqrt{S(v)}} > -2mc^2\mu_1,$$

and therefore

$$t - t_0 > 2m(\mu_2 - c^2\mu_1),$$

so that $t - t_0$ tends to $+\infty$ as m tends to $+\infty$. In case $\dot{v}_0 > 0$ the variable v tends towards s_2 monotonely; if $\dot{v}_0 < 0$ the variable v first increases monotonely, reaching the value s_1 after a finite time, and then decreases monotonely towards s_2 .

The motion approaches the side $v = s_2$ of the rectangle R as $t - t_0$ tends to $+\infty$, meeting the sides $u = r_1, u = r_2$ infinitely often and the side $v = s_1$ exactly once or not at all, depending on whether the motion starts upwards or downwards. The motion may or may not meet any one of the vertices of R on $v = s_1$, but cannot meet the vertices of R on $v = s_2$.

Case IV.— r_1, r_2, s_2 simple roots, s_1 a non-zero multiple root. The treatment and motion in this case are analogous to those in Case III.

Obviously other cases are *a priori* possible. In so far as motion in the above rectangle R is concerned, it will appear later that the above cases cover the needs of our problem.

A second type of region in which motion is possible is the infinite rectangle

$$R: u \geq r_0, \quad s_2 \leq v \leq s_1,$$

where $r_0 \geq c$ and $R(r_0) = 0$ with $R(u) > 0$ for $u > r_0$. The quantities s_1, s_2 retain their previous significance. Such a region occurs when $h \geq 0$, and we shall consider five cases of motion, the motion in each case being supposed to pass through an interior point (u_0, v_0) of R when $t = t_0$.

Case V.— r_0, s_1, s_2 simple roots. Equations (10) may be written in the form (10''), and noting that $R(u)$ is a polynomial of at most the fourth degree in u it is apparent from the second of these equations that for $t - t_0$ to tend to $+\infty$ it is necessary that u tend to $+\infty$. When u tends

to $+\infty$ the first integral in the first of the equations (10'') tends towards a finite value, since the degree of $R(u)$ is at least three, and therefore n can take only finite values. Since n can take only finite values the second equation shows that $t-t_0$ tends to $+\infty$ as u tends to $+\infty$. In case $\dot{u}_0 > 0$ the variable u tends monotonely towards $+\infty$; if $\dot{u}_0 < 0$ the variable u first decreases monotonely, reaching the value r_0 after a finite time, and then increases monotonely towards $+\infty$. This follows from Lemma 1 if the motion meets $u=r_0$ at a point not a vertex of R and from Lemmas 2 and 3 in case the motion meets a vertex of R . In case the motion meets, for example, the vertex (r_0, s_2) of R and this vertex is the point $(c, -c)$, the motion in the (u, v) -plane corresponds to a collision in the (x, y) -plane of the material point P with the mass K.

The motion tends to the right and leaves the finite portion of the plane as $t-t_0$ tends to $+\infty$, meeting the sides $v=s_1, v=s_2$ at most a finite number of times and the side $u=r_0$ exactly once or not at all, depending on whether the motion starts towards the left or towards the right. The motion may or may not meet the vertices of R on $u=r_0$.

Case VI.— r_0 a non-zero multiple root, s_1, s_2 simple roots. If $\dot{u}_0 > 0$, the treatment is the same as in Case V. When $\dot{u}_0 < 0$, the value of u tends in monotone decreasing fashion towards r_0 as $t-t_0$ increases. From the first equation in (10'') it follows that n must tend to $+\infty$ as u tends to r_0 . An argument similar to that employed in Case II. will now show that $t-t_0$ tends to $+\infty$ as u tends to r_0 .

The motion moves monotonely towards the right as in Case V., or else the motion moves monotonely towards the left, touching both $v=s_1, v=s_2$ infinitely often and approaching the side $u=r_0$ as $t-t_0$ tends to $+\infty$, depending on whether the motion starts towards the right or towards the left. The motion meets neither of the vertices of R .

Case VII.— r_0, s_1 simple roots, s_2 a non-zero multiple root. Here we employ equations (10). In order that $t-t_0$ tend to $+\infty$ it is necessary that u tend to $+\infty$. Now

$$\begin{aligned} t-t_0 &> \int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} - c^2 \int_{v_0}^v \frac{dv}{\sqrt{S(v)}} \\ &= \int_{u_0}^u \frac{u^2 du}{\sqrt{R(u)}} - c^2 \int_{u_0}^u \frac{du}{\sqrt{R(u)}} = \int_{u_0}^u \frac{u^2 - c^2}{\sqrt{R(u)}} du, \end{aligned}$$

and therefore when u tends to $+\infty$ so will $t-t_0$. Since the first integral in the first equation of (10) tends towards a finite limit as u tends to $+\infty$ the second integral in this equation must do likewise. Therefore the variable v cannot tend to the value s_2 , since s_2 is a multiple root of $S(v)=0$. Accordingly v approaches a limit $v^*(s_2 < v^* \leq s_1)$ as $t-t_0$ tends to $+\infty$.

In order to determine whether v ever attains the value s_1 consider the functions

$$F(u_0) = \int_{u_0}^{+\infty} \frac{du}{\sqrt{R(u)}}, \quad G(v_0) = \int_{v_0}^{s_1} \frac{dv}{\sqrt{S(v)}},$$

in which $\sqrt{R(u)}$, $\sqrt{S(v)}$ have the signs of \dot{u} , \dot{v} respectively. $G(v_0)$ decreases monotonely and continuously from $+\infty$ to 0 as v_0 ranges from s_2 to s_1 . Clearly $\dot{v}_0 > 0$ is necessary for v to take the value s_1 , and for simplicity let us suppose $\dot{u}_0 > 0$. Suppose u_0 to have a fixed value u_0^* . $F(u_0^*)$ will have a fixed finite value, and there exists exactly one value v_0^* of v_0 such that

$$G(v_0) \gtrless F(u_0^*) \quad \text{according as} \quad v_0 \lessgtr v_0^*.$$

If $v_0 > v_0^*$, the variable v takes the value $v=s_1$, otherwise it does not, with v tending towards s_1 if $v=v_0^*$ or towards a value v^* between s_2 and s_1 if $v_0 < v_0^*$ as $t-t_0$ tends to $+\infty$.

The above analysis holds without essential modification if $\dot{u}_0 < 0$.

The motion tends to the right and leaves the finite part of the plane, approaching the line $v=v^(s_2 < v^* \leq s_1)$ as $t-t_0$ tends to $+\infty$. The motion never meets the side $v=s_2$, meets the side $v=s_1$ at most once, and meets the side $u=r_0$ exactly once or not at all according as the motion starts to the left or to the right. The motion may or may not meet the vertex of R on $v=s_1$.*

Case VIII.— r_0 , s_2 simple roots, s_1 a non-zero multiple root. This case is analogous to the preceding one.

Case IX.— r_0 a simple root, s_1 , s_2 non-zero multiple roots. Using the methods employed in Case VII. it is seen that the necessary and sufficient condition for $t-t_0$ to tend to $+\infty$ is that u tend to $+\infty$. The variable v approaches a limit $v^*(s_2 < v^* < s_1)$ and never takes the values $v=s_1$, $v=s_2$.

The motion tends to the right and leaves the finite part of the plane, approaching a line $v=v^(s_2 < v^* < s_1)$ as $t-t_0$ tends to $+\infty$. The motion never meets the sides $v=s_1$, $v=s_2$ and meets the side $u=r_0$ exactly once or not at all according as the motion starts to the left or to the right.*

The above cases of motion comprise all that will be needed for motion in a non-degenerate rectangle, and we now take up the case where the rectangle degenerates into a line segment.

LEMMA 4.—*In order that motion occur on $u=u_0$ ($u_0 > c$) it is necessary that $u=u_0$ be a multiple root of the equation $L(u)=0$. The corresponding necessary condition for motion on $v=v_0$ ($|v_0| < c$) is that $v=v_0$ be a multiple root of the equation $M(v)=0$.*

The proof of this lemma is immediate in view of (6), (7), and (8).

Case X.— u_0 a multiple root, v_1, v_2 simple roots. Setting $u=u_0$ in (9), and integrating the second equation, we replace (9) and (10) by

$$\dot{v} = \frac{\sqrt{S(v)}}{u_0^2 - v^2}, \quad t - t_0 = \int_{v_0}^v \frac{u_0^2 - v^2}{\sqrt{S(v)}} dv$$

respectively.

The motion is periodic and moves back-and-forth from one endpoint to the other on the straight line segment $u=u_0, v_2 \leq v \leq v_1$.

Case XI.— u_1, u_2 simple roots, v_0 a multiple root. This case is analogous to Case X.

Case XII.— u_0 a simple root with $R(u) > 0$ for $u > u_0$, v_0 a multiple root. A slight modification of the procedure in Case X. shows that the motion is as follows :—

Motion occurs on the line $v=v_0$ and tends to the right, leaving the finite part of the plane as $t-t_0$ tends to $+\infty$. The motion meets the line $u=u_0$ exactly once or not at all according as the motion starts to the left or to the right.

PART II.—UNEQUAL MASSES ($0 < K' < K$).

3. *Classification of the roots of $R(u)=0, S(v)=0$ for $h \neq 0$.*—The nature of the motion is determined by the roots of $R(u)=0, S(v)=0$. In preparation for the classification of the forms of motion we shall now classify the roots of the equations

$$\begin{aligned} R(u) &= 2(u^2 - c^2)(hu^2 + (K + K')u + k) = 0, \\ S(v) &= 2(v^2 - c^2)(hv^2 + (K - K')v + k) = 0, \end{aligned}$$

in which we regard the integration constants h, k as varying parameters. Both equations possess the roots $\pm c$ for all values of h and k , and we consider instead the equations

$$\begin{aligned} L(u) &= hu^2 + (K + K')u + k = 0, \\ M(v) &= hv^2 + (K - K')v + k = 0. \end{aligned}$$

Let u_1, u_2 denote the roots of $L(u)=0$, and v_1, v_2 denote the roots of $M(v)=0$.

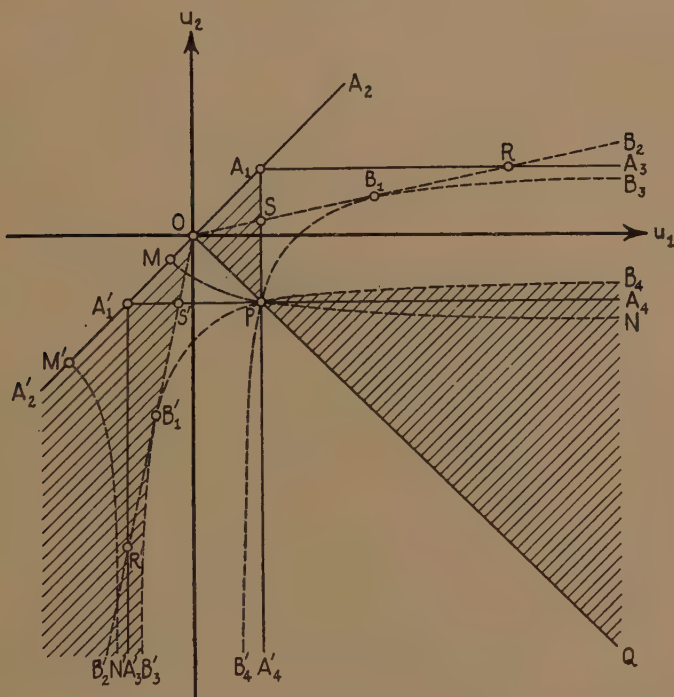
Since we assume $h \neq 0$, each equation always has two roots. If the roots u_1, u_2 are complex, so are v_1, v_2 . Suppose u_1, u_2 are real, the notation being taken so that $u_2 \leq u_1$. If now u_1, u_2 are regarded as the rectangular coordinates of a point in a (u_1, u_2) -plane, the totality of real roots of $L(u)=0$ is represented by the points of the half-plane $u_2 \leq u_1$, the points on the line $u_1 + u_2 = 0$ being excluded, since $K + K' \neq 0$. The roots of $L(u)=0$ are represented by the points of the upper quadrant $u_2 \leq u_1, u_1 + u_2 > 0$, when $h < 0$, and by the points of the lower quadrant $u_2 \leq u_1, u_1 + u_2 < 0$ when $h > 0$.

Once u_1, u_2 are known v_1, v_2 may be computed by means of the equations

$$\mathbf{T}: v_1+v_2=\sigma(u_1+u_2), \quad v_1v_2=u_1u_2, \quad (\sigma=(K-K')/(K+K')),$$

in which the notation is taken so that $v_2 \leq v_1$, when v_1 and v_2 are real. We view T as a point transformation applied to the points of the above two quadrants. The coordinates of a point before transformation are the roots of $L(u)=0$ for certain fixed values of h and k ; the coordinates of the transformed point are the roots of $M(v)=0$ for the same values of h and k . The problem of classifying the roots of these equations is therefore reduced to a study of the transformation of the half-plane $u_2 \leq u_1$ by T .

Fig. 1.



In the above figure $A_2'A_2$ represents the line $u_2=u_1$ and OQ the line $u_1+u_2=0$, every point of which is a fixed point under T . The lines A_1A_3 , $A_1'A_4$ represent $u_2=c$, $u_2=-c$ respectively; the lines A_1A_4' , $A_1'A_3'$ represent $u_1=c$, $u_1=-c$ respectively. The figure has been labelled so that the transforms of the points $B_1, B_2, B_3, B_4, B_1', B_2', B_3', B_4'$ by T are the points $A_1, A_2, A_3, A_4, A_1', A_2', A_3', A_4'$ respectively. The two rays OB_2, OB_2' have

$$\frac{u_2}{u_1} = \frac{1 - (1 - \sigma^2)^{1/2}}{1 + (1 - \sigma^2)^{1/2}}, u_1 \geq 0; \quad \frac{u_2}{u_1} = \frac{1 + (1 - \sigma^2)^{1/2}}{1 - (1 - \sigma^2)^{1/2}}, u_1 \leq 0$$

respectively for equations, and are respectively tangent at B_1, B_1' to the branches of the hyperbolas

$$u_1u_2 - c\sigma(u_1+u_2) + c^2 = 0, \quad u_1u_2 + c\sigma(u_1+u_2) + c^2 = 0, \quad . \quad . \quad (14)$$

represented by $B_4'PB_1B_3, B_4PB_1'B_3'$. The transformation T carries the points of the regions $A_2OB_2, A_2'OB_2'$ inclusive of OA_2 and OA_2' , but exclusive of OB_2 and OB_2' , into points with complex coordinates. The remaining part of the half-plane $u_2 \leq u_1$, which is bounded by the broken line $B_2'OB_2$, is carried by T into the half-plane $v_2 \leq v_1$, the rays OB_2, OB_2' being transformed into the rays OA_2, OA_2' respectively. The regions $OPB_1O, OPB_1'O$ are transformed into the triangles $OPA_1O, OPA_1'O$ respectively, the boundaries OB_1, OB_1', PB_1, PB_1' going over into OA_1, OA_1', PA_1, PA_1' respectively. The regions $B_3B_1PB_4, B_3'B_1'PB_4'$ are transformed into the infinite rectangles $A_3A_1PA_4, A_3'A_1'PA_4'$ respectively, the boundaries $B_1B_3, B_1'B_3', PB_4, PB_4'$ going over into $A_1A_3, A_1'A_3', PA_4, PA_4'$ respectively. The regions $B_3B_1B_2, B_3'B_1'B_2'$ are transformed into the regions $A_3A_1A_2, A_3'A_1'A_2'$ respectively, the boundaries $B_1B_2, B_1'B_2'$ going over into $A_1A_2, A_1'A_2'$ respectively. Finally the regions $B_4PQ, B_4'PQ$ are transformed into the regions $A_4PQ, A_4'PQ$ respectively.

We have already noted that the line OQ is to be excluded. In addition to this the shaded regions, together with the sides $OA_1, PA_1, PB_4, OA_2', B_1'B_3'$, may be excluded on the following grounds: the region OPA_1O , together with the sides OA_1, PA_1 (A_1 included), are excluded because in all cases $u_2 \leq u_1 \leq c$, and therefore, since $h < 0$, it follows from (8) that $R(u) < 0$ for $u > c$. From (9) it follows that motion with $u > c$ cannot occur. The region B_4PQ and the side PB_4 are eliminated, since for these points $v_1 > c, v_2 \leq -c$. Since $h < 0$, it follows from (8) that $S(v) < 0$ for $|v| < c$. Reference to (9) then shows that motion with $|v| < c$ is impossible. The region $B_2'OA_2'$ (including OA_2' , but excluding OB_2') is eliminated, since for these points v_1, v_2 are complex numbers, and, therefore, inasmuch as $h > 0$, it follows from (8) that $S(v) < 0$ for $|v| < c$. The region $B_3'B_1'B_2'$ (including its entire boundary) is excluded, for in this case $v_2 \leq v_1 \leq -c$, and therefore from (8), since $h > 0$, it follows that $S(v) < 0$ for $|v| < c$.

In addition it is possible to rule out the case where $L(u) = 0$ has complex roots. We have already pointed out that in this case the roots of $M(v) = 0$ are also complex. When $h < 0$ we have $R(u) < 0$ for $u > c$, and motion with $u > c$ is therefore impossible from (9). If $h > 0$, we have $S(v) < 0$ for $|v| < c$, and motion with $|v| < c$ is consequently impossible from (9).

The remaining portions of the half-plane $u_2 \leq u_1$ give rise to motions and are listed in the tables of the next section. The arcs MN and $M'N'$ are employed in the discussion of motion of the x -axis, and will be disregarded until § 5.

TABLE I.

	Position.	Roots.					Rectangle of motion.
		$L(u)=0.$	$M(v)=0.$	$r_2.$	$r_1.$	$\delta_2.$	$\delta_1.$
1.....	A_1A_2	$c < u_2 = u_1$	$v_1 \neq \bar{v}_1 = v_2$	u_1^\dagger	u_1^\dagger	$-c$	c
2.....	$A_2A_1RB_2$	$c < u_2 < u_1$	$v_1 \neq \bar{v}_1 = v_2$	u_2	u_1	$-c$	$u_2 \leq u \leq u_1, -c \leq v \leq c$
3.....	RB_2	$c < u_2 < u_1$	$c < v_2 = v_1$	u_2	u_1	$-c$	$u_2 \leq u \leq u_1, -c \leq v \leq c$
4.....	B_2RA_3	$c < u_2 < u_1$	$c < v_2 < v_1$	u_2	u_1	$-c$	$u_2 \leq u \leq u_1, -c \leq v \leq c$
5.....	A_1R	$c = u_2 < u_1$	$v_1 \neq \bar{v}_1 = v_2$	c^\dagger	u_1	$-c$	$c < u \leq u_1, -c \leq v \leq c$
6.....	R	$c = u_2 < u_1$	$c < v_2 = v_1$	c^\dagger	u_1	$-c$	$c < u \leq u_1, -c \leq v \leq c$
7.....	RA_3	$c = u_2 < u_1$	$c < v_2 < v_1$	c^\dagger	u_1	$-c$	$c < u \leq u_1, -c \leq v \leq c$
8.....	RA_1SB_1R	$u_2 < c < u_1$	$v_1 \neq \bar{v}_1 = v_2$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v \leq c$
9.....	B_1R	$u_2 < c < u_1$	$c < v_2 = v_1$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v \leq c$
10.....	$A_3RB_1B_3$	$u_2 < c < u_1$	$c < v_2 < v_1$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v \leq c$
11.....	SB_1	$u_2 < c < u_1$	$0 < v_2 = v_1 < c$	c	u_1	$\left\{ \begin{array}{l} -c \\ v_1^\dagger \\ v_1^\dagger \end{array} \right\}$	$\left\{ \begin{array}{l} -c \leq v < v_1 \\ v = v_1 \\ v_1 < v \leq c \end{array} \right\}$
12.....	B_1	$u_2 < c < u_1$	$v_2 = v_1 = c$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v < c$
13.....	B_1B_3	$u_2 < c < u_1$	$c = v_2 < v_1$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v < c$
14.....	B_1SFB_1	$u_2 < c < u_1$	$-c < v_2 < v_1 < c$	c	u_1	$\left\{ \begin{array}{l} -c \\ v_1 \end{array} \right\}$	$\left\{ \begin{array}{l} -c \leq v \leq v_2 \\ v_1 \leq v \leq c \end{array} \right\}$
15.....	PB_1	$u_2 < c < u_1$	$-c < v_2 < v_1 = c$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v \leq v_2$
16.....	$B_3B_1PB_4$	$u_2 < c < u_1$	$-c < v_2 < c < v_1$	c	u_1	$-c$	$c \leq u \leq u_1, -c \leq v \leq v_2$

† Multiple root.

4. *Classification of Motions.*—As a first basis for classification we group the motions into three classes according as $h \leq 0$. These classes will then be further subdivided into types.

Class I. $h < 0$.—The motions in this class are represented by the points in the non-excluded portions of the upper quadrant. These points are classified in Table I., together with the roots of $L(u)=0$, $M(v)=0$, $R(u)=0$, $S(v)=0$, and the corresponding rectangles of motion in the (u, v) -plane. The first column of the table refers to the position of the points in the quadrant; the various regions (without boundaries) are listed with three or more letters, the boundaries of the regions (without endpoints) by two letters, and the endpoints of the boundaries by single letters.

The motions in Class I. will now be resolved into types, the numbers in the parentheses below referring to Table I.

TYPE A (1).—Motion occurs on the ellipse $u=u_1$. Cf Case X. in § 2.

TYPE B (2, 3, 4).—Motion takes place in the annular region bounded by the ellipses $u=u_1$, $u=u_2$ meeting these ellipses and the segments of the x -axis connecting them infinitely often. Cf. Case I. in § 2.

TYPE C (5, 6, 7).—Motion occurs within the ellipse $u=u_1$. It meets this ellipse at most once and tends towards the segment KK' of the x -axis as $t-t_0$ tends to $+\infty$, each of the segments of the x -axis connecting K and K' to the ellipse $u=u_1$ being met infinitely often. Collisions with the masses K and K' do not occur. Cf. Case II. in § 2.

TYPE D (8, 9, 10).—Motion occurs within the ellipse $u=u_1$. It meets this ellipse and the segment of the x -axis joining K to K' as well as each of the segments connecting K and K' to the ellipse $u=u_1$ infinitely often. Collisions with the masses K and K' may or may not occur. Cf. Case I. in § 2.

TYPE E (11).—Motion occurs within the ellipse $u=u_1$, meeting this ellipse and the segment KK' of the x -axis infinitely often, and may take three forms. When $-c \leq v < v_1$ ($v_1 < v \leq c$) the motion meets the segment of the x -axis connecting $K(K')$ to the ellipse $u=u_1$ at most once or possibly collides with $K(K')$, and then tends through increasing (decreasing) v -values towards the segment of the hyperbola $v=v_1$ within the ellipse as $t-t_0$ tends to $+\infty$. For $v=v_1$ motion of a pendulum-like character occurs on the above arc of $v=v_1$. Cf. Cases III. and IV. in § 2 for the first two forms of motion and Case XI. for the third.

TYPE F (12, 13).—Motion occurs within the ellipse $u=u_1$, meeting this ellipse and the segment KK' of the x -axis infinitely often. The segment

of the x -axis connecting K to the ellipse $u=u_1$ is met at most once with a possible collision at K, and the motion tends through increasing v -values towards the segment of the x -axis connecting K' to the ellipse when $t-t_0$ tends to $+\infty$. Cf. Case IV. in § 2.

TYPE G (14).—Motion occurs within the ellipse $u=u_1$, either in the region $-c \leq v \leq v_2$ about K or in the region $v_1 \leq v \leq c$ about K'. Collision with K in the first case or with K' in the second case may or may not occur. Cf. Case I. in § 2.

TYPE H (15, 16).—Motion occurs within the ellipse $u=u_1$ in the region $-c \leq v \leq v_2$ about K. Collision with K may or may not occur. Cf. Case I. in § 2.

Class II. $h > 0$.—The motions in this class are represented by the points in the non-excluded portions of the lower quadrant. These points are classified in Table II.

The motions in Class II. may now be divided into types, the numbers in the parentheses below referring to Table II.

TYPE A (1, 2, 3).—Motion occurs on the hyperbola $v=v_1$. The motion meets the segment KK' of the x -axis at most once and leaves the finite part of the plane as $t-t_0$ tends to $+\infty$. Cf. Case XII. in § 2.

TYPE B (4, 5, 6).—Motion occurs in the region $v_2 \leq v \leq v_1$. The motion meets the segment KK' of the x -axis at most once, the hyperbolas $v=v_1$, $v=v_2$ at most a finite number of times, and leaves the finite part of the plane as $t-t_0$ tends to $+\infty$. Cf. Case V. in § 2.

TYPE C (7).—Motion occurs in the region $-c < v \leq v_1$. The motion meets the segment KK' of the x -axis and the hyperbola $v=v_1$ at most once, leaving the finite part of the plane with $\lim v=v^* (-c < v^* \leq v_1)$ as $t-t_0$ tends to $+\infty$. The motion never collides with K. Cf. Case VII. in § 2.

TYPE D (8).—Motion occurs in the region $-c \leq v \leq v_1$, meeting the segment KK' of the x -axis at most once. The motion meets the hyperbola $v=v_1$ and the x -axis to the right of K at most a finite number of times, leaving the finite part of the plane as $t-t_0$ tends to $+\infty$. The motion may or may not collide with K. Cf. Case V. in § 2.

TYPE E (9).—Motion occurs in the region $-c \leq v < c$, meeting the segment KK' and the part of the x -axis to the right of K at most once. The motion leaves the finite part of the plane with $\lim v=v^* (-c \leq v^* < c)$ as $t-t_0$ tends to $+\infty$. The motion may or may not collide with K. Cf. Case VIII. in § 2.

TABLE II.

	Position.	Roots.				Rectangle of motion.
		$L(u)=0.$	$M(v)=0.$	$r_0.$	$s_2.$	$s_1.$
1.....	OS'	$-c < u_2 < u_1 < c$	$-c < v_2 = v_1 < 0$	c	v_1^\dagger	$u \geq c, v = v_1$
2.....	S'	$-c = u_2 < u_1 < c$	$-c < v_2 = v_1 < 0$	c	v_1^\dagger	$u \geq c, v = v_1$
3.....	S'B ₁ '	$u_2 < -c < u_1 < c$	$-c < v_2 = v_1 < 0$	c	v_1^\dagger	$u \geq c, v = v_1$
4.....	OS'PO	$-c < u_2 < u_1 < c$	$-c < v_2 < v_1 < c$	c	v_2	$u \geq c, v_2 \leq v \leq v_1$
5.....	S'P	$-c = u_2 < u_1 < c$	$-c < v_2 < v_1 < c$	c	v_2	$u \geq c, v_2 \leq v \leq v_1$
6.....	B ₁ 'PS'B ₁ '	$u_2 < -c < u_1 < c$	$-c < v_2 < v_1 < c$	c	v_2	$u \geq c, v_2 \leq v \leq v_1$
7.....	PB ₁ '	$u_2 < -c < u_1 < c$	$-c = v_2 < v_1 < c$	c	$-c^\dagger$	$u \geq c, v_2 \leq v \leq v_1$
8.....	B ₁ 'PB ₁ 'B ₂ '	$u_2 < -c < u_1 < c$	$v_2 < -c < v_1 < c$	c	$-c$	$u \geq c, -c < v \leq v_1$
9.....	PB ₂ '	$u_2 < -c < u_1 < c$	$v_2 < -c < v_1 = c$	c	$-c$	$u \geq c, -c \leq v < c$
10.....	A ₁ 'PB ₂ '	$u_2 < -c < v_1 < c$	$v_2 < -c < c < v_1$	c	$-c$	$u \geq c, -c \leq v \leq c$
11.....	PA ₂ '	$u_2 < -c < c = u_1$	$v_2 < -c < c < v_1$	c^\dagger	$-c$	$u > c, -c \leq v \leq c$
12.....	QPA ₂ '	$u_2 < -c < c < u_1$	$v_2 < -c < c < v_1$	u_1	$-c$	$u \geq u_1, -c \leq v \leq c$

† Multiple root.

TYPE F (10).—The motion meets the segment KK' of the x -axis at most once and each of the segments of the x -axis exterior to KK' at most a finite number of times, leaving the finite part of the planes as $t-t_0$ tends to $+\infty$. The motion may or may not collide with either one of K or K' . Cf. Case V. in § 2.

TYPE G (11).—Motion occurs in the region $u > c$. The motion leaves the finite part of the plane meeting the segments of the x -axis exterior to KK' a finite number of times or else the motion tends to the segment KK' of the x -axis, meeting the segments of the x -axis exterior to KK' infinitely often as $t-t_0$ tends to $+\infty$. Cf. Case VI. in § 2.

TYPE H (12).—Motion occurs exterior to the ellipse $u=u_1$ and meets this ellipse at most once. The motion leaves the finite part of the plane meeting the segments of the x -axis exterior to the ellipse $u=u_1$ at most a finite number of times as $t-t_0$ tends to $+\infty$. Cf. Case V. in § 2.

TABLE III.

	Roots.					Rectangle of motion.
	$L(u)=0.$	$M(v)=0.$	$r_0.$	$s_2.$	$s_1.$	
1.....	$u_1 < c$	$-c < v_1 < c$	c	$-c$	v_1	$u \geq c, \quad -c \leq v \leq v_1$
2.....	$u_1 < c$	$v_1 = c$	c	$-c$	c^\dagger	$u \geq c, \quad -c \leq v < c$
3.....	$u_1 < c$	$v_1 > c$	c	$-c$	c	$u \geq c, \quad -c \leq v \leq c$
4.....	$u_1 = c$	$v_1 > c$	c^\dagger	$-c$	c	$u > c, \quad -c \leq v \leq c$
5.....	$u_1 > c$	$v_1 > c$	u_1	$-c$	c	$u \geq u_1, \quad -c \leq v \leq c$

Class III. $h=0$.—Equations (8) yield

$$L(u)=(K+K')u+k=0, \quad M(v)=(K-K')v+k=0. \quad (15)$$

Let $u=u_1$, $v=v_1$ respectively be the roots of these equations. We regard u_1 , v_1 as rectangular coordinates of a point (u_1, v_1) in a (u_1, v_1) -plane, and it is readily seen that the various pairs of roots of these equations may be represented by the points of the line

$$(K+K')u_1=(K-K')v_1$$

drawn in the (u_1, v_1) -plane. This geometrical representation makes apparent the cases listed in Table III.

The remaining possibility $u_1 < c$, $v_1 \leq -c$ may be ruled out, because we should then have $S(v) < 0$ for $|v| < c$ from (8), and therefore from (9) motion with $|v| < c$ would be impossible.

† Multiple root.

A comparison of Table III. with Table II. shows that the motions for $h=0$ fall under Types D, E, F, G, H of Class II.

5. *Motion on the x-axis.*—In §1 mention was made of the possibility of motion on the x -axis. We shall now investigate this motion. The differential equations of motion take three forms, namely :—

$$\left. \begin{aligned} (x^2-c^2)\dot{x}^2 &= 2(hx^2 + (K+K')x + (K-K')c - hc^2) = U(x), \\ (x^2-c^2)\dot{x}^2 &= 2(hx^2 - (K-K')x - (K+K')c - hc^2) = V(x), \\ (x^2-c^2)\dot{x}^2 &= 2(hx^2 - (K+K')x - (K-K')c - hc^2) = W(x), \end{aligned} \right\} \quad (16)$$

according as $x > c$, $|x| < c$, $x < -c$ respectively. Integration yields

$$\begin{aligned} t-t_0 &= \int_{x_0}^x \sqrt{\frac{x^2-c^2}{U(x)}} dx, \quad t-t_0 = \int_{x_0}^x \sqrt{\frac{x^2-c^2}{V(x)}} dx, \\ t-t_0 &= \int_{x_0}^x \sqrt{\frac{x^2-c^2}{W(x)}} dx, \quad . \quad . \quad . \quad (17) \end{aligned}$$

so that $t=t_0$ when $x=x_0$. We shall begin with a discussion of the roots of $U(x)=0$, $V(x)=0$, $W(x)=0$ for values of $h \neq 0$. Denote by u_1 , u_2 the roots of $U(x)=0$, by v_1 , v_2 the roots of $V(x)=0$, by w_1 , w_2 the roots of $W(x)=0$. We have the following relations between roots and coefficients :—

$$\begin{aligned} h(u_1+u_2) &= -K-K', & hu_1u_2 &= (K-K')c - hc^2, \\ h(v_1+v_2) &= K-K', & hv_1v_2 &= -(K+K')c - hc^2, \\ h(w_1+w_2) &= K+K', & hw_1w_2 &= -(K-K')c - hc^2. \end{aligned}$$

Substituting the values of h obtained from the equations of the first column into equations of the second column, the equations become

$$h(u_1+u_2) = -K-K', \quad u_1u_2 + c\sigma(u_1+u_2) + c^2 = 0, \quad . \quad . \quad (18)$$

$$h(v_1+v_2) = K-K', \quad v_1v_2 + c\sigma^{-1}(v_1+v_2) + c^2 = 0, \quad . \quad . \quad (19)$$

$$h(w_1+w_2) = K+K', \quad w_1w_2 + c\sigma(w_1+w_2) + c^2 = 0, \quad . \quad . \quad (20)$$

Referring to (14) and fig.1, it is apparent that the roots of $U(x)=0$, $W(x)=0$ are given by the coordinates of the points where the branch of the hyperbola $B_4PB_1'B_3'$ intersects the lines

$$h(u_1+u_2) = -(K+K'), \quad h(u_1+u_2) = K+K'$$

respectively †.

The roots of $V(x)=0$ are given by the coordinates of the points where the line

$$h(u_1+u_2) = K-K' \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

† The results on the roots of $L(u)=0$, $M(v)=0$ obtained in §5 may, of course, be obtained in an analytical manner. For the sake of uniformity of treatment we have retained the geometrical approach employed previously.

intersects the arcs of the hyperbola

$$u_1 u_2 + c\sigma^{-1}(u_1 + u_2) + c^2 = 0, \quad (22)$$

represented † in fig. 1 by MN and M'N'.

Class I. $h < 0$.—We divide the motions into three types according as $x > c$, $x < -c$, $|x| < c$.

TYPE α . $x > c$.—Here the roots u_1, u_2 of $U(x) = 0$ are given by the coordinates of the points on the arc PB_4 (P excluded). We accordingly have $-c < u_2 < c < u_1$, so that $U(x) > 0$ for $c \leq x < u_1$. It follows readily from (16) and (17) that the motion is as follows :

If the initial velocity is positive the motion proceeds to the right, and after a finite time reaches $x = u_1$; thereupon it proceeds to the left, and ends in a finite time with a collision at K. If the initial velocity is negative, motion is to the left, and ends after a finite time with collision at K.

TYPE β . $|x| < c$.—Let

$$h_1 = -\frac{(\sqrt{K} - \sqrt{K'})^2}{2c}, \quad h_2 = -\frac{(\sqrt{K} + \sqrt{K'})^2}{2c}.$$

It may now be verified that if h lies in the interval $h_2 < h < h_1$ the line (21) meets the hyperbola (22) at no real point, and therefore $V(x) = 0$ has no real roots. If $h = h_1$ the line is tangent to the hyperbola at the point M' and $V(x) = 0$ has a double root given by

$$v_1 = v_2 = -c \frac{\sqrt{K} + \sqrt{K'}}{\sqrt{K} - \sqrt{K'}} < -c.$$

If $h = h_2$ the line is tangent to the hyperbola at the point M and $V(x) = 0$ has a double root given by

$$-c < v_1 = v_2 = -c \frac{\sqrt{K} - \sqrt{K'}}{\sqrt{K} + \sqrt{K'}} < 0. \quad (23)$$

If $h_1 < h < 0$ the line intersects the arc M'N' and $V(x) = 0$ has two real roots for which

$$v_2 < v_1 < -c.$$

If $h < h_2$ the line intersects the arc MP and $V(x) = 0$ has two real roots for which

$$-c < v_2 < v_1 < c. \quad (24)$$

Consequently, if $h_2 < h < 0$ the inequality $V(x) < 0$ holds throughout the interval $|x| \leq c$. If $h = h_2$ the inequality $V(x) < 0$ holds for $|x| \leq c$ except at the point (23), where $V(x) = 0$ has a double root. For $h < h_2$ the function $V(x)$ has simple zeros at two points v_1, v_2 of the interval $|x| < c$.

† It is interesting to note that the hyperbolic arcs MN, M'N' are the transforms by T of the line segments S'A₄, R'A₃' respectively.

Between these points $V(x) > 0$, with $V(x) < 0$ at the remaining points of the interval.

The nature of the motion between K and K' may now easily be determined from (16) and (17), and is as follows :—

(i.) If $h_2 < h < 0$ the motion proceeds to the right or to the left on KK', and ends after a finite time by collision with K or K', depending on the sign of the initial velocity.

(ii.) If $h = h_2$ an equilibrium motion $x \equiv v_1$, where v_1 is given in (23), is possible. If the initial position lies to the left of the equilibrium point and the initial velocity is positive the motion proceeds constantly towards the right, approaching the equilibrium point as $t - t_0$ tends to $+\infty$. When the initial velocity is negative the motion proceeds constantly to the left, ending with a collision at K' after a finite time. Analogous results hold when the initial position of the motion lies to the right of the equilibrium point.

(iii.) If $h < h_2$ motion occurs in the interval $-c < x \leq v_2$ ($v_1 \leq x < c$). The motion tends to the right (left) if the initial velocity is positive (negative), and reaches $x = v_2$ ($x = v_1$) after a finite time; thereupon it proceeds to the left (right), and ends in a finite time with a collision at K' (K). If the initial velocity is negative (positive) motion is to the left (right), and ends after a finite time with a collision at K' (K).

TYPE γ . $x < -c$.—The roots w_1, w_2 of $W(x) = 0$ are given by the coordinates of the points on the arc PB_3' (P excluded). We have $u_2 < -c < u_1 < c$, so that $W(x) > 0$ for $u_2 < x \leq -c$. The motion is analogous to that in Type α .

Class II. $h > 0$.—The motions are divided into three types as in Class I.

TYPE α . $x > c$.—The roots of $U(x) = 0$ are given by the coordinates of the points on the arc PB_3' (P excluded). We accordingly have $u_2 < -c < u_1 < c$, so that $U(x) > 0$ for $x \geq c$. From (16) and (17) it follows that the motion is as follows :—

The motion tends to the right and leaves the finite part of the x -axis as $t - t_0$ tends to $+\infty$ if the initial velocity is positive; otherwise it moves to the left, colliding with K after a finite time.

TYPE β . $|x| < c$.—The roots $V(x) = 0$ are given by the coordinates of the points on the arc PN (P excluded). It follows that $v_2 < -c, c < v_1$, so that $V(x) < 0$ for $|x| \leq c$.

The motion collides with one of the masses after a finite time, striking K or K' according as the initial velocity is positive or negative.

TYPE γ . $x < -c$.—The roots of $W(x) = 0$ are given by the coordinates of the points on the arc PB_4 (P excluded). We accordingly have $-c < w_2 < c < w_1$, so that $W(x) > 0$ for $x \leq -c$. The motion is of the same nature as in Type α above.

Class III. $h=0$.—Setting $h=0$ in (16), it is easy to verify that

$$U(x)>0, x\geq c; V(x)<0, |x|\leq c; W(x)>0, x\leq -c.$$

The motions in this class may accordingly be divided into three types identical with α, β, γ in Class II.

6. *Critique of Charlier's Classification.*—From a study of Charlier's work † it appears that he has not observed that once the roots of one of the equations $L(u)=0, M(v)=0$ are given the roots of the other are completely determined. The neglect of this fact led him to introduce various forms of motion in his classification which do not exist. Thus, to cite one example, the form of motion classified by Charlier as IV B β on p. 136, which in our notation becomes

$$u_1>u_2=c, \quad -c<v_2\leq v_1<c,$$

does not exist. Referring to fig. 1, the roots u_1, u_2 of $L(u)=0$ would be given by the coordinates of a point (u_1, u_2) on the line A_1A_3 , and the roots v_1, v_2 of $M(v)=0$ by the coordinates of a point (v_1, v_2) in the triangle $A_1A_1'P$. Now it is easy to verify that T carries no point of the line A_1A_3 into a point of the triangle $A_1A_1'P$, and the non-existence is thereby demonstrated.

This same oversight is responsible for Charlier's omission of the motion classified in the present paper as Type E in Class II. In place of this Charlier considers a case VL on p. 139, which in our notation becomes

$$v_1=c>v_2.$$

For $h>0$ Charlier assumes $|v_2|<c$, which is not permissible. That $h>0, v_1=c$ is incompatible with $|v_2|<c$ may be seen from fig. 1, in which $h>0, v_1=c$ is represented by the line PA'_4 (which is the transform of PB'_4 by T), and therefore $v_2<-c$ necessarily.

On p. 130 Charlier assumes that a monotone increasing function $u_1=u_1(t)$ (λ in Charlier's notation) defined for all $t>t_0$ necessarily tends to $+\infty$ as t tends to $+\infty$. This is, of course, an unwarranted assumption, and in § 2 above we have established the correctness of this assumption by considering the differential equation satisfied by $u_1(t)$. In this connexion it may not be amiss to point out that Charlier apparently was under a misapprehension as to the behaviour of the motion in the neighbourhood of the point at infinity. Thus, for example, in Case II. α δ on p. 131 he implies that the motion in Class II., Type B above touches each of the hyperbolas $v=v_1, v=v_2$ infinitely often as $t-t_0$ tends to $+\infty$, whereas, as we have pointed out, it touches them only a finite number of times.

† Charlier, *op. cit.* The references to Charlier in § 6 refer to this work.

For the treatment of rectilinear motion Charlier employs the differential equations (9) by setting in them $u \equiv c$ or $v \equiv \pm c$. As pointed out in § 1, this procedure is open to objection. It may be seen, however, that the differential equations resulting are in agreement with those obtained by us in § 5, provided k (α in Charlier's notation) be assigned suitable values †. This was not done by Charlier, and extraneous cases of motion are again introduced. Thus, for example, Case I. $\alpha\gamma$ on p. 125 may be shown to be impossible, since $V(-c)$, $V(c)$ are both negative, as may be seen from (16).

PART III.—EQUAL MASSES ($K' = K$).

7. *Classification of Motions for $h \neq 0$.*—The transformation T now becomes

$$T: v_1 + v_2 = 0, \quad v_1 v_2 = u_1 u_2,$$

so that the half-plane $u_2 \leq u_1$ is transformed into the line $v_1 + v_2 = 0$, $v_1 \geq 0$. Referring to fig. 1, rays OB_2 , OB_2' in the figure become the positive u_1 -axis and the negative u_2 -axis respectively; the hyperbolic branches $B_4'PB_1B_3$, $B_4PB_1'B_3'$ coalesce into a single branch

$$u_1 u_2 + c^2 = 0, \quad u_1 > 0,$$

the points B_1 and B_1' going off to infinity as indicated in fig. 2.

For the present we exclude motion on the x -axis. As in Part II., we rule out the case where $L(u) = 0$ has complex roots. The shaded regions and the sides OQ , OA_1 , OA_2' , PA_1 , PB_1 of fig. 2 may be neglected. For $h < 0$ the portions A_1A_2 , A_2A_1R , A_1R , RA_1SB_1 , SB_1 , B_1SPB_1 of fig. 2 correspond respectively to motions of Types A, B, C, D, E, G in Class I., Types F and H no longer existing. For $h > 0$ the portions OB_1' , $B_1'POB_1'$, $A_4'PB_1'$, PA_4' , PQA_4' of fig. 2 correspond respectively to the motions of Types A, B, F, G, H in Class II. On the line PB_1' we have

$$u_2 < -c < u_1 < c, \quad v_2 = -c, \quad v_1 = c,$$

so that motion in the (u, v) -plane takes place in the rectangle

$$u \geq c, \quad -c < v < c,$$

a form of motion coming under Type E in Class II., as may be seen on comparing Cases VII., VIII., and IX. in § 2. Accordingly motions of Types C, D together with certain forms in Type E no longer exist.

8. *Classification of motions for $h = 0$.*—The functions $L(u)$, $M(v)$ are now given by

$$L(u) = 2Ku + k, \quad M(v) = k,$$

and consequently from (8) and (9) motion with $|v| < c$ requires that $k \leq 0$. If u_1 denotes the root of $L(u) = 0$, the necessary and sufficient condition for $k \leq 0$ is that $u_1 \geq 0$. When $u_1 = k = 0$ motion occurs on

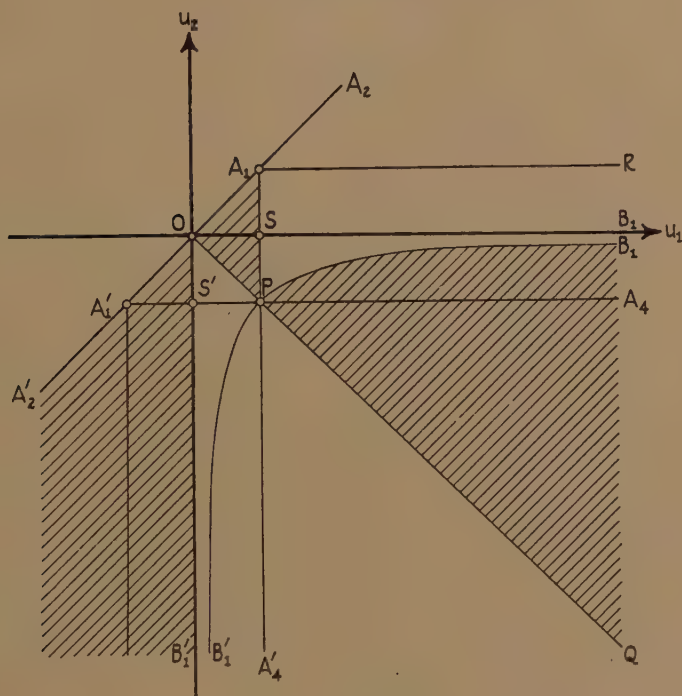
† Cf. (16) of the present paper.

$v=\text{const.}$, a form of motion not present for $h=0$ when $K'>K$. The remaining cases $0<u_1<c$, $u_1=c$, $u_1>c$ lead to motions of the same type as 3, 4, 5 respectively in Table III. for Class III.

9. *Motion on the x -axis.*—Setting $K=K'$ in (18) and (20), it appears that the branch of the hyperbola $B_4PB_1'B_3'$ in fig. 1 becomes the branch $B_1'PB_1$ in fig. 2, the roots of $U(x)=0$, $W(x)=0$ being given now by the coordinates of the points where $B_1'PB_1$ intersects the lines

$$h(u_1+u_2)=-2K, \quad h(u_1+u_2)=2K$$

Fig. 2.



respectively. The roots of $V(x)=0$ are given by

$$v_1 = \sqrt{\frac{2Kc}{h} + c^2}, \quad v_2 = -\sqrt{\frac{2Kc}{h} + c^2}.$$

It may now be verified that the character of the motion on the x -axis for $K'=K$ is identical with that for $K'<K$. As is to be expected, greater symmetry prevails when $K'=K$. Thus when $K'=K$ we have $v_1=0$ in (ii.) and $v_1=-v_2$ in (iii.) of Type β of Class I. in § 5.

University of Maryland,
College Park, Maryland, U.S.A.

XVI. *On the Critical Dimensions of Tuned Transmitting
Circular Loop Aerials.*

By S. S. BANERJEE, D.Sc., Department of Physics,
Benares Hindu University *.

[Received July 26, 1938.]

Introduction.

WITH the advent of ultra short waves in radio communications the use of loop aerials has been considerably increased for transmitting as well as for receiving purposes. These loop aerials have the double advantage of possessing comparatively smaller dimensions and also a sharp directive property, both of which are very essential for communications with mobile wireless stations. When the wave-lengths employed for communications are very small the dimensions of the loop become comparable with the wave-lengths, and under such conditions Palmer and his collaborators^{(1), (2), (3), (4)} have recently shown that rectangular loop aerials possess certain critical dimensions for transmission and reception of maximum energy which are quite independent of the usual conditions of tuning. Their method of finding out the condition for maximum loop current consists in calculating the vector sum of the "direct" and "indirect" currents in each side of the loop, which is considered to be an oscillating electric doublet. In the present investigation we have determined mathematically the critical dimensions of tuned transmitting circular loop aerials for obtaining the maximum current in them by "retarded induced potential" method. This method is particularly preferred for the reason that the induced effect of every part of the aerial wire on the rest of the whole circumference of the same may be taken into account. It may be mentioned that this method can also be extended to the rectangular and other types of loop aerials. It has got the added superiority of practically uniform radiation of energy in all the directions in the plane of the loop owing to the absence of any sharp bend along the perimeter. The above dimensions for radiation of maximum energy from the circular loop transmitting aerial have also been verified experimentally, and it is shown that the dimensions obtained are very critical for availing the maximum transmission of energy from the aerial.

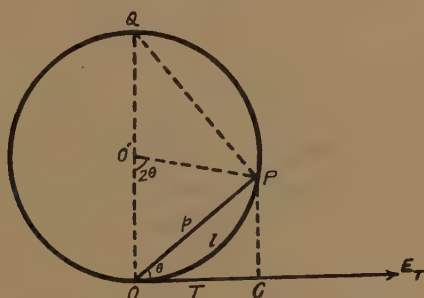
* Communicated by the Author.

Theoretical.

The theoretical method, as has been mentioned before, consists in calculating the critical dimensions of a tuned transmitting circular loop aerial for flowing of maximum current through it. Now, in order to find out the above dimensions we will obtain the condition for maximum induced electric intensity along the whole length of the aerial wire. For this, it will suffice if we deduce the condition for maximum induced electric intensity at any given point of the wire, due to the passage of high frequency current in the rest of the wire.

Let OPQ (fig. 1) represent the circular transmitting loop aerial and suppose a high frequency sinusoidal current, represented by the real part of $Ie^{j\omega t}$, is entering at O and passing along the length of the aerial wire, where I , j , ω and t have their usual significance. It will be evident that, as it is a circular aerial, a given element of this wire is situated exactly in the same manner as any other element with respect to the remaining portion of the aerial.

Fig. 1.



The electric intensity at any point O (fig. 1) along the length of the wire is given by the relation

$$E_T = -\text{grad}_T \psi - \frac{1}{c} \frac{dA}{dt}, \quad \dots \dots \dots (1)$$

where ψ and A are the scalar and vector potentials respectively at O, due to the high frequency current in the whole length of the aerial, and c is the velocity of light. The direction of E_T coincides with the tangent at O, and T is the distance from O to the foot of the perpendicular dropped from any point on the circumference of the circle to the tangent at O.

Now, if P be any point on the wire at a distance p from O, l the length OP of the wire, and L the total length of the wire, then the values of A and ψ are given by

$$A = \frac{q}{c} \int_0^L \frac{[i]}{p} dl \quad \dots \dots \dots (2)$$

After substituting the values of $\text{grad} \psi$ and $\frac{dA}{dt}$ so obtained in equation (1), we get

$$\begin{aligned} E_T = & -j \frac{2I}{c} e^{j\omega t} \left[\frac{1}{D \cos 2\theta \sin \theta} \{ \cos (mD \sin \theta) \cos (mD\theta) \right. \\ & \left. - j \sin (mD \sin \theta) \cos (mD\theta) \} \right]_0^{\pi/2} \\ & - \frac{2I}{c^2} \omega e^{j\omega t} \int_0^{\pi/2} \frac{1}{\sin \theta} [\sin (mD\theta) \sin (mD \sin \theta) \\ & + j \cos (mD \sin \theta) \sin (mD\theta)] d\theta. \end{aligned} \quad (10)$$

Neglecting the imaginary part in the above equation, we get

$$\begin{aligned} E_T = & K \left[\frac{\sin (mD \sin \theta) \cos (mD \theta)}{D \cos 2\theta \sin \theta} \right]_0^{\pi/2} \\ & - Km \int_0^{\pi/2} \frac{\sin (mD \sin \theta) \sin (mD\theta)}{\sin \theta} d\theta, \end{aligned} \quad (11)$$

where

$$K = \frac{2I}{c} e^{j\omega t}.$$

The integral in the second expression of the right-hand side of equation (11) cannot be put in any suitable form. Therefore, the term $\sin (mD \sin \theta)$ in the expression is expanded, and an approximate value of E_T is obtained by neglecting the terms containing higher powers of mD , on the assumption that mD is not very high, which is also found in practice.

Thus from equation (11) we may write

$$\begin{aligned} E_T = & Km \left[\frac{\cos (mD\theta)}{\cos 2\theta} - \frac{m^2 D^2 \sin^2 \theta \cos (mD\theta)}{6 \cos 2\theta} \right]_0^{\pi/2} \\ & - Km \int_0^{\pi/2} mD \sin (mD\theta) d\theta \\ = & Km \left[\frac{m^2 D^2}{6} \cos \frac{mD\pi}{2} - \cos \frac{mD\pi}{2} - 1 \right] - Km \left[1 - \cos \frac{mD\pi}{2} \right], \\ = & K \frac{m^3 D^2}{6} \cos \frac{mD\pi}{2} - 2Km. \end{aligned} \quad (12)$$

In order to obtain the condition for evaluating maximum value of E_T , it is differentiated with respect to D and equated to zero, and thus we get

$$\frac{dE_T}{dD} = \frac{Km^3}{6} \left[2D \cos \frac{mD\pi}{2} - \frac{mD^2\pi}{2} \sin \frac{mD\pi}{2} \right] = 0. \quad (13)$$

From this we find that either

$$D = 0 \quad (14)$$

or
$$2 \cos \frac{mD\pi}{2} - \frac{mD\pi}{2} \sin \frac{mD\pi}{2} = 0. \quad . \quad . \quad . \quad . \quad . \quad (15)$$

The value of $D=0$ does not require any consideration as it gives the minimum value of E_T . This may be examined by substituting the value of $D=0$ in the expression for $\frac{d^2E_T}{dD^2}$ obtained from equation (13) and given by

$$\frac{d^2E_T}{dD^2} = \frac{Km^3}{6} \left[2 \cos \frac{mD\pi}{2} - mD\pi \sin \frac{mD\pi}{2} - \frac{m^2D^2\pi^2}{4} \cos \frac{mD\pi}{2} \right], \quad (16)$$

which yields a positive value of $\frac{d^2E_T}{dD^2}$.

Now, from equation (15) we get the condition for obtaining the maximum value of E_T , given by

$$\frac{mD\pi}{2} \tan \frac{mD\pi}{2} = 2,$$

or, $mD\pi = 2.16, 7.39$, etc.

From these values of $mD\pi$ we find that the critical dimensions for the circumference of the circular loop aerial for maximum current will be given by $D\pi = 0.35 \lambda, 1.17 \lambda$, etc., where λ is the wave-length. For practical purposes, however, we will not consider the higher values of the circumference, as then the value of mD will also become fairly high. It may be mentioned that the above values of $mD\pi$ yield negative values of d^2E_T/dD^2 , which shows that these solutions correspond to maximum values of E_T .

Experimental Arrangement and Observations.

The experimental arrangement made to verify the results obtained in the previous section is shown in fig. 2. L represents the tank inductance of an ultra-short wave oscillator emitting waves of about 2 to 4 metres. C is the tuning variable condenser of low value (maximum capacity $0.000018 \mu F$). A is the circular loop aerial connected through a radio-frequency thermo-galvanometer T , to measure the current flowing through the aerial. The oscillator was carefully shielded from all sides in order to avoid the stray-capacity effects.

Circular loop aeriels of different sizes, ranging from 50 cm. to 4.1 metres in circumference, were introduced across the inductance L , and in each case the maximum current in the thermo-galvanometer T was noted by adjusting the variable condenser C . At the same time it was also noted that the reading of variable condenser for maximum current in the aerial of different dimensions was not appreciably altered. This is a necessary condition to be maintained by not allowing the dimensions of the aeriels

to become so small as to make its inductance comparable with the tank inductance L .

In order to test whether the aerials of critical dimensions were actually radiating maximum energy a set of horizontal dipole aerial was connected

Fig. 2.



TABLE I.

Wave-length (λ) = 2.61 metres.

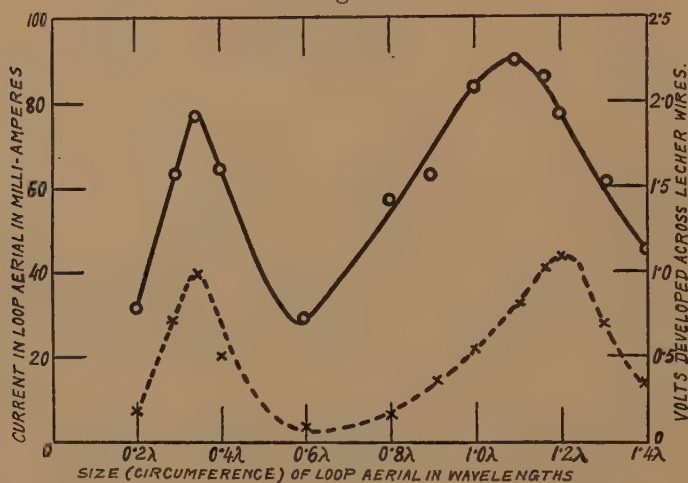
Size (circumference) of the circular loop aerial.	Current in the circular loop aerial in milliamperes.	Voltage developed across the Lecher wires, in volts.
0.20 λ	32.5	0.20
0.30 λ	64.4	0.75
0.35 λ	78.0	0.98
0.40 λ	64.6	0.50
0.60 λ	30.2	0.10
0.80 λ	58.3	0.18
0.90 λ	62.0	0.33
1.00 λ	83.5	0.57
1.10 λ	90.0	0.85
1.17 λ	85.6	1.05
1.20 λ	78.0	1.10
1.30 λ	60.5	0.72
1.40 λ	47.0	0.35

to the two ends of a pair of Lecher wires, and the voltage developed at the first potential antinode was measured with a Moullin's Thermionic Voltmeter. This system of the horizontal dipole was kept at a distance of several wave-lengths away from the ultra-short wave oscillator and the arrangement was found to be quite sensitive.

The observations recorded in order to verify the theoretical relation of the critical dimensions of the loop aerial are shown in Table I. It will be observed from the second column of the table that the currents in the loop aerial attain maximum values for the dimensions of the circumference as 0.35λ and 1.10λ , which agree well with the theoretically calculated critical dimensions of the circular aerial shown in the previous section.

Fig. 3 indicates graphically in continuous line the variation of the current in the circular loop aerial as its length of the perimeter is altered. The radiated voltage developed across the Lecher wires with the horizontal dipole for different sizes of the loop are also shown in the same figure with dotted lines. It will be observed from fig. 3 that the maximum

Fig. 3.



radiated voltage for the first critical dimension of the loop aerial, viz., 0.35λ as the circumference, agrees well with the maximum current produced in it. But for the second critical dimension of the loop, viz., 1.17λ as the circumference, the corresponding maxima do not agree very closely. This may, however, be due to the fact that the value of mD in this case does not remain very small, as has been assumed in deriving the theoretical conditions in the last section. The observations were repeated for several wave-lengths to verify the above critical dimensions for the circular loop aerals, and similar results have been obtained.

Summary and Conclusions.

Critical dimensions for maximum radiation of energy from circular loop transmitting aerals, the sizes of which are comparable with the wave-length transmitted, have been derived mathematically by the

“retarded induced potential” method. It has been found that for transmission of maximum energy the circumference of the circular loop aerial should be 0.35λ , 1.17λ , etc. The first two critical dimensions have been verified experimentally with waves of 2 to 4 metres length. It may be concluded from the above results that special attention should be directed to the dimensions of the ultra-short wave circular loop aerials for availing the maximum energy for transmission.

I desire to express my sincere thanks to Prof. P. Dutt, M.A.(Cantab), Head of the Department of Physics, for his helpful suggestions during the progress of the present work. Lastly, I have great pleasure in thanking my colleague, Mr. B. N. Singh, M.Sc., for his help, which he offered while carrying out the experimental part of the investigation.

References.

- (1) Palmer, L. S., Proc. Roy. Soc. A, cxxxvi. p. 193 (1932).
- (2) Palmer, L. S., and Honeyball, L. L. K., Proc. Inst. Rad. Eng. xx. p. 1345 (1932).
- (3) Palmer, L. S., and Taylor, D., Proc. Phys. Soc. xlv. p. 62 (1934).
- (4) Palmer, L. S., and Taylor, D., Proc. Inst. Rad. Eng. xxii. p. 93 (1934).

Physics Laboratory,
College of Science,
Benares Hindu University,
Benares (India).

XVII. *On the Problem of Wave-motion for Sub-Infinite Domains.*

By DR. ARNOLD N. LOWAN †.

[Received July 26, 1938.]

IN a previous paper ‡ the author discussed the problem of integrating the general differential equation of wave-motion

$$(I.) \nabla^2 \mathbf{U}(P, t) = 2b \frac{\partial}{\partial t} \mathbf{U}(P, t) + \frac{1}{a^2} \frac{\partial^2}{\partial t^2} \mathbf{U}(P, t) + \Phi(P, t)$$

for the infinite domain, of one, two, or three dimensions, the solution satisfying the initial conditions

$$(II.) \lim_{t \rightarrow 0} \mathbf{U}(P, t) = f(P),$$

$$(III.) \lim_{t \rightarrow 0} \frac{\partial}{\partial t} \mathbf{U}(P, t) = g(P).$$

The object of the present paper is the integration of the system I., II., III., for the domains and the boundary conditions given in the following tables :—

A.

	Range of variables.	Boundary conditions.
1.....	$0 < x < \infty$	$\phi(P, t) = \phi(t) \quad \text{for } x=0$
2.....	$\left\{ \begin{array}{l} 0 < x < \infty \\ -\infty < y < \infty \end{array} \right\}$	$\phi(P, t) = \phi(y) \quad \text{for } x=0$
3.....	$\left\{ \begin{array}{l} 0 < x < \infty \\ -\infty < y < \infty \\ -\infty < z < \infty \end{array} \right\}$	$\phi(P, t) = \phi(y, z) \quad \text{for } x=0$

† Communicated by the Author.

‡ "On Wave-motion for Infinite Domains" (Phil. Mag. xxvi. p. 340 (1938)). This paper will be referred to as "A.N.L."

B.

	Range of variables.	Boundary conditions.
1.....	—	—
2.....	$\left\{ \begin{array}{l} 0 < x < \infty \\ 0 < y < \infty \end{array} \right\}$	$\phi(P, t) = \phi_1(y)$ for $x=0$ $\phi(P, t) = \phi_2(x)$ for $y=0$
3.....	$\left\{ \begin{array}{l} 0 < x < \infty \\ 0 < y < \infty \\ -\infty < z < \infty \end{array} \right\}$	$\phi(P, t) = \phi_1(y, z)$ for $x=0$ $\phi(P, t) = \phi_2(x, z)$ for $y=0$

C.

	Range of variables.	Boundary conditions.
1.....	—	—
2.....	—	—
3.....	$\left\{ \begin{array}{l} 0 < x < \infty \\ 0 < y < \infty \\ 0 < z < \infty \end{array} \right\}$	$\phi(P, t) = \phi_1(y, z)$ for $x=0$ $\phi(P, t) = \phi_2(x, z)$ for $y=0$ $\phi(P, t) = \phi_3(x, z)$ for $z=0$

In the absence of a better term the above domains, bounded in one direction but extending to infinity in one, two, or three directions, are referred to as sub-infinite domains.

PART I.—DOMAIN A.

In order to obtain the solution for the domain A we make the substitution

$$U(P, t) = u(P, t) + v(P, t), \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where $u(P, t)$ satisfies the system

$$D \left\{ \begin{array}{l} \nabla^2 u(P, t) = 2b \frac{\partial}{\partial t} u(P, t) + \frac{1}{a^2} \frac{\partial^2}{\partial t^2} u(P, t) + \Phi(P, t) \quad 0 < x < \infty, \quad (2) \\ \lim_{t \rightarrow 0} u(P, t) = f(P) \quad 0 < x < \infty, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3) \\ \lim_{t \rightarrow 0} \frac{\partial}{\partial t} u(P, t) = g(P) \quad 0 < x < \infty, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4) \\ u(P) = 0 \text{ for } x=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5) \end{array} \right.$$

and $v(P, t)$ satisfies the system

$$\text{E} \left\{ \begin{array}{l} \nabla^2 v(P, t) = 2b \frac{\partial}{\partial t} v(P, t) + \frac{1}{a^2} \frac{\partial^2}{\partial t^2} u(P, t) \quad 0 < x < \infty, \quad \dots \quad (6) \\ \lim_{t \rightarrow 0} v(P, t) = 0, \quad \dots \quad (7) \\ \lim_{t \rightarrow 0} \frac{\partial}{\partial t} v(P, t) = 0, \quad \dots \quad (8) \\ v(P, t) = \phi(P, t) \text{ for } x=0, \quad \dots \quad (9) \end{array} \right.$$

where

$$\left. \begin{array}{l} \phi(P, t) = \phi(t) \quad \text{for the one-dimensional case,} \\ \phi(P, t) = \phi(y, t) \quad \text{for the two-dimensional case,} \\ \phi(P, t) = \phi(y, z, t) \quad \text{for the three-dimensional case.} \end{array} \right\} \quad \dots \quad (10)$$

The first three equations of the system D are identical with those treated in A.N.L., except that they now refer to the domain $x > 0$.

The solution of our present system D may be obtained from the corresponding solutions treated in A.N.L. by assuming that the semi-infinite domain is extended from $x=0$ to $x=-\infty$, and that for $x < 0$ the function $f(P)$ is defined by the relations

$$\left. \begin{array}{l} f(-x) = -f(x) \quad \text{for the one-dimensional case,} \\ f(-x, y) = -f(x, y) \quad \text{for the two-dimensional case,} \\ f(-x, y, z) = -f(x, y, z) \quad \text{for the three-dimensional case,} \end{array} \right\} \quad \dots \quad (11)$$

and similar relations for the functions $g(P)$ and $\Phi(P, t)$.

An analysis of the solutions obtained in the four sections of A.N.L. leads to the conclusion that if the relations (11) are satisfied for each of the three functions, $f(P)$, $g(P)$, and $\Phi(P, t)$, these solutions satisfy the boundary condition (5). Our problem thus reduces to the integration of the system E.

If, as in Section II. of A.N.L., we make the substitutions

$$v(x, t) = e^{kt} \cdot v_1(x, t); \quad \phi_1(P, t) = e^{kt} \cdot \phi(P, t), \quad \dots \quad (12)$$

where $k = a^2 b$, the Laplace transform $v_1^*(P, p)$ of the functions $v_1(P, t)$ must satisfy the differential equation

$$\nabla^2 v_1^*(P, p) - \frac{1}{a^2} (p^2 - k^2) v_1^*(P, p) = 0 \quad \dots \quad (13)$$

and the boundary condition

$$v_1(P, p) = \phi_1^*(P, p) \text{ for } x=0, \quad \dots \quad (14)$$

where $\phi_1^*(P, p)$ is the Laplace transform of $\phi_1(P, t)$ defined in (10) and (12).

We proceed to solve the systems (13) and (14) for the case of one, two, or three dimensions.

SECTION I.—*The One-dimensional Case.*

In this case the function $v_1^*(x, p)$ must satisfy the differential equation

$$\frac{d^2}{dx^2} v_1^*(x, p) - \frac{1}{a^2} (p^2 - k^2) v_1^*(x, p) = 0 \quad . \quad . \quad . \quad (15)$$

and the boundary condition

$$v_1^*(0, p) = \phi_1^*(p), \quad . \quad . \quad . \quad . \quad . \quad (16)$$

whence

$$v_1^*(x, p) = \phi_1^*(p) e^{-\frac{x}{a} \sqrt{p^2 - k^2}} \quad . \quad . \quad . \quad . \quad . \quad (17)$$

In order to obtain the inverse Laplace transform of (17) we proceed as follows :—

Starting with the identity—(equation (13) A.N.L.)

$$\frac{e^{-\eta \sqrt{p^2 - k^2}}}{\sqrt{p^2 - k^2}} = \int_{\eta}^{\infty} e^{-pt} J_0(k \sqrt{\eta^2 - t^2}) dt \quad . \quad . \quad . \quad (18)$$

we obtain by differentiation with respect to η

$$e^{-\eta \sqrt{p^2 - k^2}} = e^{-p\eta} + k\eta \int_{\eta}^{\infty} e^{-pt} \frac{J_1(k \sqrt{\eta^2 - \tau^2})}{\sqrt{\eta^2 - \tau^2}} d\tau \quad . \quad . \quad . \quad (19)$$

In view of (19) (17) becomes

$$v_1^*(x, p) = \phi_1^*(p) e^{-p \frac{x}{a}} + k \frac{x}{a} \int_{\frac{x}{a}}^{\infty} \frac{J_0(k \sqrt{(x/a)^2 - \tau^2})}{\sqrt{(x/a)^2 - \tau^2}} \cdot e^{-p\tau} \phi_1^*(p) d\tau \quad (20)$$

But, according to a well-known theorem †, if

$$L\{\phi_1(t)\} = \phi_1^*(p),$$

then

$$L\{\phi_1(t - \eta)\} = \phi_1^*(p) e^{-p\eta}.$$

Accordingly the inversion of (20) yields

$$v_1(x, t) = \phi_1 \left(t - \frac{x}{a} \right) + k \frac{x}{a} \int_{\frac{x}{a}}^{\infty} \phi_1(t - \tau) \frac{J_1(k \sqrt{(x/a)^2 - \tau^2})}{\sqrt{(x/a)^2 - \tau^2}} d\tau \quad . \quad (21)$$

Thus, for the one-dimensional case, $v(x, t) = e^{ba^2t} \cdot v_1(x, t)$ where $v_1(x, t)$ is given by (21).

SECTION II.—*The Two-dimensional Case.*

In this case the function $v_1^*(x, y; p)$ must satisfy the differential equation

$$\frac{\partial^2}{\partial x^2} v_1^*(x, y; p) + \frac{\partial^2}{\partial y^2} v_1^*(x, y; p) - \frac{1}{a^2} (p^2 - k^2) v_1^*(x, y; p) = 0 \quad . \quad (22)$$

and the boundary condition

$$v_1^*(0, y; p) = \phi_1^*(y, p). \quad . \quad . \quad . \quad . \quad . \quad (23)$$

† See, for instance, John R. Carson, 'Electric Circuit Theory, and the Operational Calculus,' p. 47 (1926).

The expression $e^{-\lambda x} \cos \xi(y-\eta)$ is a solution of (22) provided that

$$\lambda^2 = \alpha^2 + \frac{1}{a^2} (p^2 - k^2).$$

Accordingly

$$v_1^*(x, y; p) = \frac{1}{\pi} \int_0^\infty d\xi \int_{-\infty}^\infty \phi_1^*(\eta, p) e^{-\frac{x}{a}} \sqrt{p^2 - (k^2 - a^2 \xi^2)} \cdot \cos \xi(y-\eta) d\eta \quad (24)$$

is a solution of (22), satisfying the boundary condition (23).

Using the results we obtained in the one-dimensional case, the inversion of (24) ultimately yields

$$\begin{aligned} v_1(x, y; t) = & \frac{1}{\pi} \int_{\xi=0}^\infty \int_{\eta=-\infty}^\infty \phi_1\left(\eta; t - \frac{x}{a}\right) \cos \xi(y-\eta) d\xi d\eta \\ & + k \frac{x}{\pi a} \int_{\xi=0}^\infty \int_{\eta=-\infty}^\infty \int_{\tau=\frac{x}{a}}^\infty \cos \xi(y-\eta) \frac{J_1\left[\sqrt{(k^2 - a^2 \xi^2)} \left(\left(\frac{x}{a}\right)^2 - \tau^2\right)\right]}{\sqrt{\left(\frac{x}{a}\right)^2 - \tau^2}} \\ & \times \phi_1(\eta, t-\tau) d\xi d\eta d\tau. \quad (25) \end{aligned}$$

Thus, in the two-dimensional case, $v(x, y; t) = e^{ba^2 t} \cdot v_1(x, y; t)$, where $v_1(x, y; t)$ is given by (25).

SECTION III.—The Three-dimensional Case.

The expression

$$\begin{aligned} v_1^*(x, y, z; p) = & \frac{1}{4\pi^2} \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \phi_1^*(\eta, \zeta; p) \cdot e^{-\lambda x} \\ & \cos [\alpha(y-\eta) + \beta(z-\zeta)] d\eta d\zeta d\alpha d\beta \quad (26) \end{aligned}$$

is a solution of

$$\nabla^2 v_1^*(x, y, z; p) - \frac{1}{a^2} (p^2 - k^2) v_1^*(x, y, z; p) = 0,$$

provided

$$\lambda^2 = \alpha^2 + \beta^2 + \frac{1}{a^2} (p^2 - k^2).$$

Furthermore, this solution satisfies the boundary condition

$$v_1^*(x, y, z; p) = \phi_1^*(y, z; p) \quad \text{for } x=0. \quad (27)$$

(See A.N.L., equation (7), Section III.)

Proceeding in the same manner as in the previous case, the inversion of (26) ultimately yields

$$v_1(x, y, z; t) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi_1\left(\eta, \zeta; t - \frac{x}{a}\right) \cos [\alpha(y-\eta) + \beta(z-\zeta)]$$

$$\begin{aligned}
& \times d\eta d\zeta d\alpha d\beta + k \frac{x}{4\pi^2 a} \int_{-\infty}^{+\infty} \int \int \cos [\alpha(y-\eta) + \beta(z-\zeta)] \\
& \times d\eta d\zeta d\alpha d\beta \int_{\tau=\frac{x}{a}}^{\infty} \frac{J_1 \left[\sqrt{[k^2 - a^2(\alpha^2 + \beta^2)] \left[\left(\frac{x}{a}\right)^2 - \tau^2 \right]} \right]}{\sqrt{\left(\frac{x}{a}\right)^2 - \tau^2}} \\
& \cdot \phi_1(\eta, \zeta, t - \tau) d\tau. \quad \dots \dots \dots (28)
\end{aligned}$$

Thus, in the three-dimensional case, $v(x, y, z; t) = e^{bat} \cdot v_1(x, y, z; t)$, where $v_1(x, y, z; t)$ is given by (27).

PART II.—DOMAIN B.

SECTION I.—The Two-dimensional Case.

In this case we make the substitution

$$U(P, t) = u(P, t) + v(P, t) + w(P, t), \quad \dots \dots \dots (1)$$

where $u(P, t)$, $v(P, t)$ and $w(P, t)$ satisfy the systems

$$\begin{cases}
\text{F} \left\{ \begin{aligned}
& \nabla^2 u(P, t) = 2b \frac{\partial}{\partial t} u(P, t) + \frac{1}{a^2} \frac{\partial^2}{\partial t^2} u(P, t) + \Phi(P, t) & 0 < x < \infty ; 0 < y < \infty, \quad \dots (2) \\
& \lim_{t \rightarrow 0} u(P, t) = f(P) & 0 < x < \infty ; 0 < y < \infty, \quad \dots (3) \\
& \lim_{t \rightarrow 0} \frac{\partial}{\partial t} u(P, t) = g(P) & 0 < x < \infty ; 0 < y < \infty, \quad \dots (4) \\
& u(P, t) = 0 \quad \text{for } x = 0, \quad \dots \dots \dots (5) \\
& u(P, t) = 0 \quad \text{for } y = 0. \quad \dots \dots \dots (6)
\end{aligned} \right. \\
\\
\text{G} \left\{ \begin{aligned}
& \nabla^2 v(P, t) = 2b \frac{\partial}{\partial t} v(P, t) + \frac{1}{a^2} \frac{\partial^2}{\partial t^2} v(P, t) & 0 < x < \infty ; 0 < y < \infty, \quad \dots (7) \\
& \lim_{t \rightarrow 0} v(P, t) = 0 & 0 < x < \infty ; 0 < y < \infty, \quad \dots (8) \\
& \lim_{t \rightarrow 0} \frac{\partial}{\partial t} v(P, t) = 0 & 0 < x < \infty ; 0 < y < \infty, \quad \dots (9) \\
& v(P, t) = 0 \quad \text{for } x = 0, \quad \dots \dots \dots (10) \\
& v(P, t) = \phi_1(x, t) \quad \text{for } y = 0. \quad \dots \dots \dots (11)
\end{aligned} \right.
\end{cases}$$

$$\begin{cases}
 \nabla^2 w(P, t) = 2b \frac{\partial}{\partial t} w(P, t) + \frac{1}{a^2} \frac{\partial^2}{\partial t^2} v(P, t) \\
 \lim_{t \rightarrow 0} w(P, t) = 0 \\
 \lim_{t \rightarrow 0} \frac{\partial}{\partial t} w(P, t) = 0 \\
 w(P, t) = \phi_2(y, t) \text{ for } x=0, \\
 w(P, t) = 0 \text{ for } y=0.
 \end{cases}
 \begin{cases}
 0 < x < \infty ; 0 < y < \infty, \quad \dots (12) \\
 0 < x < \infty ; 0 < y < \infty, \quad \dots (13) \\
 0 < x < \infty ; 0 < y < \infty, \quad \dots (14) \\
 \dots \dots \dots (15) \\
 \dots \dots \dots (16)
 \end{cases}$$

As in Part I., the first three equations of system F are identical with those treated in A.N.L. except that they now refer to the domain $x > 0, y > 0$. The solution of the system F may be obtained from the corresponding solutions treated in A.N.L. by assuming that the domain under consideration is extended from $x=0$ to $x=-\infty$, and from $y=0$ to $y=-\infty$, and that for $x < 0$ and $y < 0$ the function $f(P)$ is defined by the relations

$$\left. \begin{aligned}
 f(-x, y) &= -f(x, y), \\
 f(x_1 - y) &= -f(x, y), \\
 f(-x_1 - y) &= f(x, y),
 \end{aligned} \right\} \dots \dots \dots (17)$$

and similar relations for the functions $g(P)$ and $\Phi(P, t)$.

An analysis of the solutions obtained in A.N.L. leads to the conclusion that if the relations (17) are satisfied for each of the three functions, $f(P)$, $g(P)$, and $\Phi(P, t)$, these solutions satisfy the boundary conditions (5) and (6). The problem thus reduces to the integration of the systems G and H.

We proceed to the consideration of the system G. The Laplace transform $v_1^*(x, y; p)$ of the function $v_1(x, y; t) = e^{-kt} v(x, t)$ must satisfy the system

$$\nabla^2 v_1^*(P, p) - \frac{1}{a^2} (p^2 - k^2) v_1^*(P, p) = 0, \quad \dots \dots \dots (18)$$

$$v_1^*(P, p) = \phi_{1,0}^*(y, t) \text{ for } x=0, \quad \dots \dots \dots (19)$$

$$v_1^*(P, p) = 0 \text{ for } y=0, \quad \dots \dots \dots (20)$$

where $\phi_{1,0}^*$ is the Laplace transform of the function

$$\phi_{1,0}(y, t) = e^{kt} \phi_1(y, t).$$

The expression

$$v_1^*(x, y; p) = \frac{2}{\pi} \int_0^\infty \int_0^\infty \phi_{1,0}^*(\eta, p) e^{-\frac{x}{a} \sqrt{p^2 - (k^2 - a^2 \alpha^2)}} \sin \alpha \eta \sin \alpha y \, d\alpha \, d\eta. \quad (21)$$

is a solution of (18), satisfying the boundary conditions (19) and (20).

From the results of Part II., Section II., the inverse Laplace transform of (21) may be written down at once in the form

$$\begin{aligned}
 v_1(x, y; t) = & \frac{2}{\pi} \int_{\alpha=0}^{\infty} \int_{\eta=0}^{\infty} \phi_{1,0} \left(\eta, t - \frac{x}{a} \right) \cdot \sin \alpha \eta \sin \alpha y d\alpha d\eta \\
 & + \frac{2kx}{\pi a} \int_{\alpha=0}^{\infty} \int_{\eta=0}^{\infty} \int_{\tau=\frac{x}{a}}^{\infty} \phi_{1,0}(\eta t - \tau) \sin \alpha \eta \sin \alpha y \\
 & \times \frac{J_1 \left[\sqrt{(k^2 - a^2 \alpha^2) \left(\left(\frac{x}{a} \right)^2 - \tau^2 \right)} \right]}{\sqrt{\left(\frac{x}{a} \right)^2 - \tau^2}} d\alpha d\eta d\tau. \quad \dots (22)
 \end{aligned}$$

A comparison between the systems G and H shows clearly that the solution of the latter system may be obtained from the solution of the former by interchanging x and y in the second member of (22) and replacing the subscript 1, in ϕ_1 by 2. Thus :

$$\begin{aligned}
 w(x, y; t) = & \frac{2}{\pi} \int_{\alpha=0}^{\infty} \int_{\xi=0}^{\infty} \phi_{2,0} \left(\xi, t - \frac{y}{a} \right) \cdot \sin \alpha \xi \sin \alpha x d\alpha d\xi \\
 & + \frac{2ky}{\pi a} \int_{\alpha=0}^{\infty} \int_{\xi=0}^{\infty} \int_{\tau=\frac{y}{a}}^{\infty} \phi_{2,0}(\xi, t - \tau) \sin \alpha x \sin \alpha \xi \\
 & \times \frac{J_1 \left[\sqrt{(k^2 - a^2 \alpha^2) \left(\left(\frac{y}{a} \right)^2 - \tau^2 \right)} \right]}{\sqrt{\left(\frac{y}{a} \right)^2 - \tau^2}} d\alpha d\xi d\tau. \quad \dots (23)
 \end{aligned}$$

SECTION II.—The Three-dimensional Case.

As in the case of Section I., we make the substitution $U(P, t) = u(P, t) + v(P, t) + w(P, t)$, where $u(P, t)$, $v(P, t)$, $w(P, t)$ satisfy systems entirely similar to F, G, and H, except that they refer to the three-dimensional case and that the functions $\phi_1(y, t)$ and $\phi_2(x, t)$ are replaced by $\phi_1(y, z; t)$ and $\phi_2(x, z; t)$.

$$Lx \phi_{1,0}(y, z, t) = e^{kt} \phi_1(y, z, t) \quad \text{and} \quad \phi_{2,0}(x, z, t) = e^{kt} \phi_2(x, z, t).$$

As before, the solution $u(P, t)$ may be obtained from the appropriate solution in A.N.L. by extending the definition of the functions $f(x, y, z)$, $g(x, y, z)$, and $\phi(x, y, z; t)$ by the relations

$$\left. \begin{aligned} f(-x, y, z) &= -f(x, y, z), \\ f(x, -y, z) &= -f(x, y, z), \\ f(-x, -y, z) &= f(x, y, z), \end{aligned} \right\}, \quad \dots \dots \dots (24)$$

and by similar relations for the functions $g(x, y, z)$ and $\Phi(x, y, z; t)$. The solution of A.N.L. is then seen to satisfy the appropriate boundary condition of the system F.

We proceed to derive the solution for the appropriate system G for the domain under consideration. The Laplace transform $v_1^*(x, y, z; p)$ of the function $v_1(x, y, z; t) = e^{-kt}v(x, y, z; t)$ must satisfy the system

$$\nabla^2 v_1^*(x, y, z; p) - \frac{1}{a^2} (p^2 - k^2) v_1^*(x, y, z; p) = 0, \quad . \quad . \quad . \quad (25)$$

$$v_1^*(0, y, z; p) = \phi_{1,0}^*(y, z; p), \quad . \quad . \quad . \quad . \quad (26)$$

$$v_1^*(x, 0, z; p) = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

The expression

$$v_1^*(x, y, z; p) = \frac{2}{\pi^2} \int_{\eta=0}^{\infty} \int_{\zeta=-\infty}^{\infty} \int_0^{\infty} \phi_{1,0}^*(\eta, \zeta; p) \times e^{-\frac{x}{a} \sqrt{p^2 - [k^2 - a^2(\beta^2 + \gamma^2)]}} \sin \beta y \sin \beta \eta \cos \gamma(z - \zeta) d\eta d\zeta d\beta d\gamma \quad (28)$$

is a solution of (25) satisfying the boundary conditions (26) and (27) †. In view of the results obtained in Part I. the inverse Laplace transform of (28) may be obtained in the form

$$\begin{aligned} v_1(x, y, z; t) = & \frac{2}{\pi^2} \int_{\eta=0}^{\infty} \int_{\zeta=-\infty}^{\infty} \int_{\beta, \gamma=0}^{\infty} \phi_{1,0} \left(\eta, \zeta; t - \frac{x}{a} \right) \\ & \times \sin \beta y \sin \beta \eta \cos \gamma(z - \zeta) d\eta d\zeta d\beta d\gamma \\ & + \frac{2kx}{\pi^2 a} \int_{\eta=0}^{\infty} \int_{\zeta=-\infty}^{\infty} \int_{\beta, \gamma=0}^{\infty} \int_{\tau=\frac{x}{a}}^{\infty} \phi_{1,0}(\eta, \zeta, t - \tau) \\ & \times \frac{J_1 \left[\sqrt{[k^2 - a^2(\beta^2 + \gamma^2)] \left(\left(\frac{x}{a} \right)^2 - \tau^2 \right)} \right]}{\sqrt{\left(\frac{x}{a} \right)^2 - \tau^2}} \sin \beta y \sin \beta \eta \\ & \times \cos \gamma(z - \zeta) d\eta d\zeta d\beta d\gamma d\tau. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29) \end{aligned}$$

As in the previous section, it can be easily seen (by comparison of the systems G and H appropriate to the domain under consideration) that the expression for $w_1(x, y, z; t)$ may be obtained from the expression for

† See Appendix.

$v_1(x, y, z; t)$ by interchanging x and y , and by replacing $\phi_{1,0}(y, z; t)$ by $\phi_{2,0}(x, z; t)$. Thus:

$$\begin{aligned}
 w_1(x, y, z; t) = & \frac{2}{\pi^2} \int_{\xi=0}^{\infty} \int_{\zeta=-\infty}^{\infty} \overline{\int_{\alpha, \gamma=0}^{\infty}} \phi_{2,0}\left(\alpha, \zeta, t - \frac{y}{a}\right) \sin \alpha x \\
 & \times \sin \alpha \xi \cos \gamma(z - \zeta) d\xi d\zeta d\alpha d\gamma \\
 & + \frac{2ky}{\pi^2 a} \int_{\xi=0}^{\infty} \int_{\zeta=0}^{\infty} \overline{\int_{\alpha, \gamma=0}^{\infty}} \int_{\tau=\frac{y}{a}}^{\infty} \phi_{2,0}(\alpha, \zeta; t - \tau) \\
 & \times \frac{J_1 \left[\sqrt{[k^2 - a^2(\alpha^2 + \gamma^2)]} \left(\left(\frac{y}{a} \right)^2 - \tau^2 \right) \right]}{\sqrt{\left(\frac{y}{a} \right)^2 - \tau^2}} \sin \alpha x \sin \alpha \xi \\
 & \times \cos \gamma(z - \zeta) d\xi d\zeta d\alpha d\gamma d\tau. \quad \dots \dots \dots (30)
 \end{aligned}$$

PART III.—DOMAIN C.

As in the preceding sections, we make the substitution,

$$U(P, t) = u(P, t) + v(P, t) + w(P, t) + s(P, t),$$

where $u(P, t)$ and $v(P, t)$, satisfy the systems

$$\begin{aligned}
 \text{M} \left\{ \begin{aligned} \nabla^2 u(x, y, z; t) &= 2b \frac{\partial}{\partial t} u(x, y, z; t) \\ &+ \frac{1}{a^2} \frac{\partial^2}{\partial t^2} u(x, y, z; t) + \Phi(x, y, z; t) \quad x > 0, y > 0, z > 0, \quad (30) \\ \lim_{t \rightarrow 0} u(x, y, z; t) &= f(x, y, z), \quad \dots \dots \dots (31) \\ \lim_{t \rightarrow 0} \frac{\partial}{\partial t} u(x, y, z; t) &= g(x, y, z), \quad \dots \dots \dots (32) \\ u(x, y, z; t) &= 0 \quad \text{for } x=0 \quad \text{or } y=0 \quad \text{or } z=0. \quad \dots \dots \dots (33) \end{aligned} \right.
 \end{aligned}$$

$$\begin{aligned}
 \text{N} \left\{ \begin{aligned} \nabla^2 v(x, y, z; t) &= 2b \frac{\partial}{\partial t} v(x, y, z; t) \\ &+ \frac{1}{a^2} v(x, y, z; t) \quad x > 0, y > 0, z > 0, \quad (34) \\ \lim_{t \rightarrow 0} v(x, y, z; t) &= 0, \quad \dots \dots \dots (35) \\ \lim_{t \rightarrow 0} \frac{\partial}{\partial t} v(x, y, z; t) &= 0, \quad \dots \dots \dots (36) \\ v(0, y, z; t) &= \phi_1(y, z; t), \quad \dots \dots \dots (37) \\ v(x, 0, z; t) &= v(x, y, 0; t) = 0. \quad \dots \dots \dots (38) \end{aligned} \right.
 \end{aligned}$$

The functions $w(P, t)$ and $s(P, t)$ satisfy systems whose first three equations are identical with the first three equations of N, and whose boundary conditions are

$$\left. \begin{aligned} w(x, 0, z; t) &= \phi_2(x, z; t), \\ w(0, y, z; t) &= w(x, y, 0; t) = 0, \end{aligned} \right\} \dots \dots \dots (39)$$

and

$$\left. \begin{aligned} s(x, y, 0; t) &= \phi_3(x, y; t), \\ s(0, y, z; t) &= s(x, 0, z; t) = 0. \end{aligned} \right\} \dots \dots \dots (40)$$

In order to obtain the solution of the system M we extend the domain of definition of the functions $f(x, y, z)$, $g(x, y, z)$, and $\Phi(x, y, z; t)$ in the following manner:

For any of the above three functions, if we change the sign of one or all of the variables, the value of the function changes sign. If we change the sign of any two of the variables the value of the function remains unchanged. Thus, for instance,

$$\left. \begin{aligned} f(-x, y, z) &= -f(x, y, z), \\ f(-x, -y, z) &= +f(x, y, z), \\ f(-x, -y, -z) &= -f(x, y, z). \end{aligned} \right\} \dots \dots \dots (41)$$

When these conditions are satisfied an analysis of the results of A.N.L. shows that the boundary conditions of the system M are satisfied. The problem reduces to the integration of the system N for the function $v(x, y, z; t)$ and of two similar systems for the functions $w(x, y, z; t)$ and $s(x, y, z; t)$.

The Laplace transform, $v_1^*(x, y, z; p)$ of the function, $v_1(x, y, z; t) = e^{-kt}v(x, y, z; t)$ must satisfy the system

$$\nabla^2 v_1^*(x, y, z; p) - \frac{1}{a^2}(p^2 - k^2)v_1^*(x, y, z; p) = 0, \dots \dots (42)$$

$$v_1^*(0, y, z; p) = \phi_{1,0}^*(y, z; p), \dots \dots \dots (43)$$

$$v_1^*(x, 0, z; p) = v_1^*(x, y, 0; p) = 0, \dots \dots \dots (44)$$

where

$$\phi_{1,0}(y, z; t) = e^{kt}\phi_1(y, z; t). \dots \dots \dots (45)$$

The expression

$$v_1^*(x, y, z; p) = \frac{4}{\pi^2} \iiint_0^\infty e^{-\frac{x}{a}\sqrt{p^2 - [k^2 - a^2(\beta^2 + \gamma^2)]}} \times \phi_{1,0}^*(\eta, \zeta; p) \sin \beta y \sin \beta \eta \sin \gamma z \sin \gamma \zeta d\eta d\zeta d\beta d\gamma \quad (46)$$

is a solution of (42), satisfying the boundary conditions (43) and (44).

In view of the results obtained in the previous section the inversion of (46) leads to the solution

$$\begin{aligned}
 v_1(x, y, z; t) = & \frac{4}{\pi^2} \iiint_{0^-}^{\infty} \phi_{1,0} \left(\eta, \zeta; t - \frac{x}{a} \right) \sin \beta \eta \sin \beta y \\
 & \times \sin \gamma z \sin \gamma \zeta d\eta d\zeta d\beta d\gamma \\
 & + \frac{4kx}{\pi^2 a} \iiint_{0^-}^{\infty} \int_{\tau=\frac{x}{a}}^{\infty} \phi_{10}(\eta, \zeta; t - \tau) \\
 & \times \frac{J_1 \left[\sqrt{[k^2 - a^2(\beta^2 + \gamma^2)] \left(\left(\frac{x}{a} \right)^2 - \tau^2 \right)} \right]}{\sqrt{\left(\frac{x}{a} \right)^2 - \tau^2}} \sin \beta y \sin \beta \eta \sin \gamma z \\
 & \times \sin \gamma \zeta d\eta d\zeta d\beta d\gamma d\tau. \quad \dots \dots \dots (47)
 \end{aligned}$$

The solution for $w_1(x, y, z; t)$ may be obtained from the expression for $v_1(x, y, z; t)$ by interchanging x and y , and by replacing $\phi_{1,0}(x, z; t)$ by $\phi_{2,0}(y, z; t)$; similarly the solution for $s_1(x, y, z; t)$ may be obtained from $v_1(x, y, z; t)$ by interchanging x and z , and by replacing $\phi_{1,0}(y, z; t)$ by $\phi_{3,0}(x, y; t)$.

Our original system I., II., III. is thus completely solved for the domains and the appropriate boundary conditions of the table given at the beginning of this paper.

APPENDIX.

Briefly, this result may be obtained as follows:—We consider the two-dimensional problem of heat conduction in the domain $x > 0$ and $y > 0$; the boundaries $x = 0$ and $y = 0$ are being kept at zero degrees, and the initial temperature being $f(x, y, z)$. Starting with the point-source solution

$$\frac{1}{4\pi kt} e^{-\frac{(x-\xi)^2 + (y-\eta)^2}{4kt}},$$

if we associate with it the solutions

$$-\frac{1}{4\pi kt} e^{-\frac{(x+\xi)^2 + (y-\eta)^2}{4kt}},$$

$$\frac{1}{4\pi kt} e^{-\frac{(x+\xi)^2 + (y+\eta)^2}{4kt}},$$

and

$$-\frac{1}{4\pi kt} e^{-\frac{(x-\xi)^2 + (y+\eta)^2}{4kt}},$$

corresponding to a source at the point $(-\xi, -\eta)$ and two sinks at the points $(-\xi, \eta)$ and $(\xi, -\eta)$, the sum of the above four solutions, $G(x, y; \xi, \eta)$ is the Green's function for the problem of heat conduction under consideration.

As is well known, the solution of our problem is then

$$\iint f(\xi, \eta) \cdot G(x, y; \xi, \eta) d\xi d\eta.$$

Making $t=0$, the double integral reduces to $f(x, y)$. With the aid of the identity

$$\frac{1}{\pi} \int_0^\infty e^{-k\alpha^2 t} \cos \alpha (x-\xi) d\alpha = \frac{1}{2\sqrt{\pi kt}} e^{-\frac{(x-\xi)^2}{4t}}$$

the double integral reduces to (28).

Brooklyn College,
Yeshiva College.

XVIII. *Studies of the Tone Quality of Organ-Pipes.*—II. *Reed-Pipes.*

By M. MOKHTAR, M.Sc. (King's College, Newcastle-on-Tyne) *.

[Received October 31, 1938.]

[Plates IV. & V.]

Introduction.

THE dependence of the tone quality on blowing pressure in the case of flue-pipes was illustrated in a preceding paper †. The present investigation is carried out to study a similar dependability in the case of reed-pipes, and to exhibit the nature of the coupling between reed and resonator.

Apparatus.

The pipe is excited by a uniform continuous stream of air from a big reservoir fed by a pump. Special care is taken to damp out the fluctuations in the blowing pressure due to the action of the pump.

The sound produced is recorded on a photographic film fitted in a revolving camera, and by the aid of a microphone, an amplifier, and a vacuum cathode-ray oscillograph. The method of recording is described in some detail in the first paper.

The actual mechanical motion of the reed is also recorded on a similar photographic film, using a very light, small mirror stuck on the reed itself and a beam of light from a pointolite lamp.

The vibration of the air in the resonating column very near to its base is recorded also photographically by the aid of a manometric capsule, having a tiny light mirror stuck on the membrane, and a beam of light. The recording of both the mechanical motion of the reed and of the response of the manometric capsule is done simultaneously and on the same photographic plate to enable a study of the change of phase to be carried out.

The reed is a beating reed in a wind supply chamber of the type commonly employed on organs. As it vibrates it closes the entry from the supply box to the resonator once in each period, so acting as a form of escapement for the air. The reed had a natural frequency of 213 per sec. under a blowing pressure of 48 cm. of water.

* Communicated by Dr. E. G. Richardson.

† Proc. Durham Phil. Soc. ix. p. 352 (1938).

The resulting note is due to the fundamental of the reed with the fundamental or one or other of the harmonics of the pipe. As the overtones of the reed are inharmonic they are not notably produced in the note given by the pipe, because they do not form a part of the note of the column of air.

The mechanical vibrations of the reed are always simple harmonic (no. 1, Pl. IV.), although the air vibrations involve the overtones proper to the form of the air column and depend on the conditions under which the pipe is excited.

Experiments.

(A) Effect of blowing pressure on the quality of the tone produced.

If the mouthpiece containing the reed is excited when uncoupled to a resonator the character of the sound heard is usually of a raucous tone, unpleasant for the ear, especially when the pressure is high. This is due to the fact that the more sudden the discontinuity of any periodic motion, the greater the relative importance of the high upper partials into which that motion may be resolved. Thus higher pressure causes the closure of the air-duct to be more sudden, and therefore the higher upper partials become more obtrusive. This is illustrated clearly in the series of oscillograms, 2, 3, 4, 5, and 6 (Pl. IV.), which are all recorded for the same mouthpiece excited under successively increasing blowing pressures of .58, .73, .80, .93, and 3.15 cm. of water respectively and without a resonator. The gradual appearance and growth of the upper partials of the reed are easily followed in the successive oscillograms. Even the first oscillogram (no. 2, Pl. IV.) which is recorded for a very low blowing pressure gives a distinct indication of these overtones, and as we proceed to the last of the series (no. 6, Pl. IV.) the overtones completely overcome the fundamental.

On coupling the mouthpiece to a resonator having a natural frequency different from that of the reed the latter either commands or is commanded by the vibrations of the former according to the degree of coupling. This problem has been studied before by Vogel *. In our case the coupling is loose and the vibration of the reed is predominant. Nevertheless, the quality of the resultant note is affected by the presence of the resonator, in that the upper partials are less distinct and therefore the tone of the pipe is purer. This is illustrated by the series of oscillograms 7, 8, and 9 (Pl. IV.) which are recorded for the same mouthpiece of the first series when coupled to a resonator in the form of a brass cylinder 45 cm. long and 3.6 cm. in diameter (natural frequency 185.5), mounted in the usual fashion directly on the supply box. The corresponding blowing pressures

* Vogel, *Ann. d. Physik*, lxii. p. 247 (1920).

were .56, .75, and 1.03 cm. of water respectively. These are to be compared with the oscillograms 2, 3, and 5 respectively, which refer to nearly the same blowing pressures but with the reed uncoupled to a resonator. The frequency of the note heard is only slightly raised by increasing the blowing pressure, although the quality is most seriously affected by the inception and growth of the upper partials as shown in all the series of oscillograms mentioned before.

To provide a further illustration to the effect of pressure the length of the resonator is now decreased to 39.5 cm. (natural frequency 215) and the two oscillograms 10 and 11 (Pl. IV.) recorded under a blowing pressure of .70 and 1.13 cm. of water respectively. In this case the natural frequency of the resonator is very near to that of the reed, and that is why the upper partials which were clear in oscillograms 3 and 5 and partially damped in 8 and 9 are here highly damped.

(B) *The nature of the coupling between reed and resonator.*

The method of recording here is slightly different from the first one, as the vibrations of the spots of light reflected from the two tiny mirrors stuck on the reed and the membrane of the manometric capsule have to be recorded simultaneously on the same photographic plate. The spreading out of the wave-forms on a time axis is done by a rotating mirror, which sweeps the spots across the sensitive plate. The camera used is an ordinary wide aperture one. Records are taken starting from a length of the resonating air column little shorter than that corresponding to the first overtone of the column to that which would give the fundamental. The whole series is recorded under one blowing pressure of 4.8 mm. of water, this being slightly above the minimum pressure required to excite the pipe. The curves given by the manometric capsule are recorded below. Those of the reed (above) are always S.H.M. and of fundamental frequency 213.

Initially, before the mirror is rotated, the two spots of light lie vertically beneath each other. Rotation of the mirror traverses them both across the plate, so that we can examine the phase relationship from left to right between the motions by observing the lie of the peaks in a vertical direction. In the first three oscillograms 12, 13, 14 (Pl. V.) the lengths of the tube are 18.2, 19, and 20 cm. respectively. Over this range two partials are prominent in the wave-form of the system; one is in unison with the reed and the other the octave of this. (The latter requires a length of stopped pipe 18.5 cm. for resonance.) This appears to be a type of relaxation oscillation in which the forcing impulse occurs once in two periods of vibration of the air column. (The relaxing is evident in the falling amplitude of alternate peaks.) In no. 15 (22 cm. length) the octave has nearly and

in no. 16 (25 cm.) completely given way to a pure tone in unison with the reed. This state of affairs is continued at no. 17 (29 cm.).

During the whole of this series the peak of the air compression just inside the pipe has lagged more and more, up to a quarter period, behind the peak of the reed's motion (complete covering of the wind orifice), and the length of the resonating column is becoming more and more out of tune with the reed until at last each hinders the vibration of the other. Consequently the pipe is silent, although under a higher blowing pressure the reed can be forced to vibrate again. Let us now keep to our former pressure and still lengthen the pipe. At a length of 39.0 cm. the air column almost jumps into vibration, being pulled down to resound to the note of the reed. When the length reaches that corresponding to best resonance the amplitude of vibration is a maximum, as shown by no. 18 (40 cm.) (Pl. V.). A further increase of the resonator length sets it out of tune again, although it continues to give forced resonance to that of the reed, but of course, with gradually diminishing response (Pl. V. no. 19 at 42 cm. ; no. 20 at 47 cm. ; no. 21 at 49 cm.). During this stage the pressure in the air column is very slightly in advance of the reed.

Theory.

The fact that with shorter length of resonator we have a S.H.M. of the driver maintaining relaxation oscillations of the column of air at double its frequency is a sufficient indication that in this reed-pipe, at any rate, the motion is being maintained by a periodic impulse and not by continuous forcing, like a swing kept in motion by a periodic impulse from a person standing on the ground.

In ordinary resonance the driving force is a quarter period in advance of the driven body, but with an impulse lasting but a short time maintenance is most effective if the impulse of pressure occurs at the maximum compression, as Rayleigh * pointed out in reference to the singing flame. The optimum length of tube will be that for which a compression transmitted down the air column from the reed returns, after reflexion at the other end, in step with the next reed pulse. In the present instance this happens when the tube is about 18.5 cm. long (no. 12, Pl. V.). If the tube is lengthened the returning compression arrives progressively later, until it arrives too late for the reed to cope with it, and the system is reduced to silence. If the tube is further lengthened until the compression is nearly in step with the next-but-one pulse from the reed the system is again maintained. Quite a wide range of phase incompatibility is permissible without intermission of the sound. Precisely the same

* Rayleigh, 'Sound,' ii. p. 227.

laws have been found by Richardson * to govern the functioning of the singing flame.

The first overtone of the column can be excited under the same conditions, but is naturally in evidence only at the shorter lengths. It will be noticed that only the quality changes with lengthening tube (nos. 12-17, Pl. V.). There is none of the gradual lowering of pitch with length such as Vogel and others have noticed. An organ-pipe of this class has a strong metal reed unlike the thin wooden reed found on the oboe and clarinet. In the orchestral instruments the reed is more highly damped than the column of air, and, in consequence, can be pulled into synchronism with the column, whereas in this reed organ-pipe the reverse takes place. The reed maintains its own frequency and impresses it on the column of air even when the latter is considerably removed from resonance. With a flue-pipe † the edge-tone is pulled out of its natural frequency by the stronger column. One must conclude that Vogel used a soft, highly damped reed in his pipes.

Acknowledgment.

My acknowledgment and appreciation are due to Dr. E. G. Richardson, who suggested and supervised this work. Thanks are also due to Mr. E. Bawtree for lending me some of the apparatus used.

* Proc. Phys. Soc. xxxv. p. 47 (1923).

† Cf. Mokhtar, *loc. cit.*

XIX. *Hydrodynamic Forces on an Accelerated Cylinder
moving in Two-dimensions.*

By M. A. OMARA, Docteur-ès-Sciences Mathématiques, Lecturer in
Applied Mathematics, The Egyptian University, Cairo (Egypt) *.

[Received July 20, 1938.]

1. *Introduction.*

THE first real attempt to generalize the contour integral formulæ of Blasius to the accelerated cylinder seems to have been made by A. Masotti (1926) in considering the uniform rotation of a cylinder in an infinite mass of liquid at rest at infinity. He applied his results to the case of a circular cylinder and found that, in addition to the inertia force, the cylinder is subject to a centripetal force depending on the circulation of the liquid round the cylinder, generalizing thus the theorem of Kutta-Joukowski. In 1927 he dealt with the general motion of the cylinder, but the formulæ he obtained contained integrated terms and a term which could not be calculated by Cauchy's theorem.

E. Carafoli (1928) established formulæ similar to those of Masotti and calculated the resultant force and moment in the case of uniform rotation of a cylinder of any cross-section.

J. Pèrès (1936) reduced the hydrodynamic forces on a cylinder moving in any manner into a lifting force and forces of inertia due to the acceleration of the cylinder.

Rosa M. Morris (1937) obtained formulæ generalizing those of Blasius and calculated by direct integration the resultant force and moment on a cylinder whose cross-section curve can be expressed parametrically.

Some time ago the present writer, following a suggestion of Professor H. Villat, was engaged on an attempt, since completed, to generalize the exceptional cases of the theorem of Kutta-Joukowski, and was led to formulæ equivalent to those of Rosa Morris, if not a little simpler, generalizing as well the Blasius contour integral expressions, and, moreover, taking into account the possible presence of cusps on the curve of cross-section of the cylinder. The terms introduced with this respect in the expressions of the resultant force and moment have been interpreted as representing forces acting each along the tangent to the curve at the corresponding cusp, generalizing thus a result obtained (1936) in

* Communicated by the Author.

the case of rectilinear motion. By applying these considerations to a flat plate moving in any manner the resultant force was found to be always normal to the plate—a generalization of the well-known result of U. Cisotti (1927).

Also a necessary modification of the generalized Joukowski's theorem has here been stated and proved.

2. Generalization of Blasius Formulæ.

When a cylinder is moving perpendicular to its axis in an infinite mass of perfect fluid at rest at infinity, with velocity components U and V along two rectangular axes Ox , Oy fixed in a normal cross-section of the cylinder and has an angular velocity ω about the origin O , the velocity of a point (x, y) is given in the complex form

$$U - iV - i\omega(x - iy) = W - i\omega\bar{z},$$

where \bar{z} is the conjugate of $z = x + iy$.

Let u and v be the positive gradients of the velocity potential; the absolute complex velocity in the liquid and the complex potential $f = \phi + i\psi$ are in general analytic functions of z and t , satisfying the relation

$$w(z, t) = \frac{\partial f(z, t)}{\partial z}.$$

The pressure p at any point is given by

$$\frac{p}{\rho} = -\frac{\partial \phi}{\partial t} - \frac{1}{2}(u^2 + v^2) + u(U - \omega y) + v(V + \omega x),$$

where ρ is the density, supposed constant (Lamb, 'Hydrodynamics,' p. 184, 6th edition, 1932).

The components X and Y of the resultant force on the cylinder are given by

$$Y + iX = \int_c p d\bar{z}, \quad \dots \dots \dots (2.1)$$

and the resultant moment M with respect to the origin is

$$M = R \int_c zp d\bar{z}, \quad \dots \dots \dots (2.2)$$

where R denotes real part, and integration in (2.1) and (2.2) is taken round the contour c of the section of the cylinder.

These equations can be reduced to formulæ analogous to those of Blasius in the following way, first

$$(u^2 + v^2) d\bar{z} = \left(\frac{\partial f}{\partial z}\right)^2 dz - 2i \frac{\partial f}{\partial z} d\psi.$$

But at the surface of the cylinder

$$d\psi = u dy - v dx = (U - \omega y) dy - (V + \omega x) dx,$$

so that, after some reduction, we get

$$\rho d\bar{z} = -\rho \frac{\partial \phi}{\partial t} d\bar{z} - \frac{1}{2}\rho \left(\frac{\partial f}{\partial z}\right)^2 dz + \rho(W - i\omega\bar{z}) d\phi.$$

Also, $2\phi = f + \bar{f}$, and therefore,

$$2 \frac{\partial \phi}{\partial t} = \frac{\partial f}{\partial t} + \frac{\partial \bar{f}}{\partial t}, \quad \text{and} \quad 2d\phi = \left(\frac{\partial f}{\partial z} + \frac{\partial \bar{f}}{\partial \bar{z}}\right) dz.$$

Equation (2.1) can now be written

$$Y + iX = -\frac{\rho}{2} \int_c \left\{ \left(\frac{\partial f}{\partial t} + \frac{\partial \bar{f}}{\partial t}\right) d\bar{z} + \left(\frac{\partial f}{\partial z}\right)^2 dz - (W - i\omega\bar{z}) \left(\frac{\partial f}{\partial z} + \frac{\partial \bar{f}}{\partial \bar{z}}\right) dz \right\}. \quad (2.3)$$

In the same way

$$z\rho d\bar{z} = -\rho z \frac{\partial \phi}{\partial t} d\bar{z} - \frac{1}{2}\rho z \left(\frac{\partial f}{\partial z}\right)^2 dz + \rho z(W - i\omega\bar{z}) d\phi.$$

The last term can be neglected, being purely imaginary, and so we get

$$M = -\frac{\rho}{2} R \int_c z \left\{ \left(\frac{\partial f}{\partial t} + \frac{\partial \bar{f}}{\partial t}\right) d\bar{z} + \left(\frac{\partial f}{\partial z}\right)^2 dz - W \left(\frac{\partial f}{\partial z} + \frac{\partial \bar{f}}{\partial \bar{z}}\right) dz \right\}. \quad (2.4)$$

(2.3) and (2.4) are the equations aimed at. The various integrations contained here can be performed by means of Cauchy's theorem, when the complex potential of the fluid motion is known. A great simplification is obtained by transforming conformally the region outside the contour c into the exterior of a circle C of radius unity, centred at the origin of a plan ζ , where the complex potential $F(\zeta, t)$ of the corresponding fluid motion is easily obtained.

3. Complex Potential $F(z, t)$ of the Motion in the ζ -plane.

Let the transformation function be $z = G(\zeta)$ and suppose that the points at infinity in the two planes correspond to one another. $G(\zeta)$ is holomorphic outside the circle C with a single pole at infinity and can be expanded in a Laurent's series

$$z = \sum_{n=-1}^{\infty} A_n \zeta^{-n}, \quad \dots \dots \dots (3.11)$$

convergent outside and on C .

It is known that the complex potential $f(z, t)$ of the motion in the z -plane must satisfy at each instant the conditions :—

- (1) $\partial f / \partial z$ is uniform outside the contour c and vanishes at infinity ;

(2) the imaginary part of f , i. e., ψ , must reduce on c to

$$\psi = Uy - Vx - \frac{1}{2}\omega(x^2 + y^2). \quad (3.2)$$

The transformation function permits to report the preceding value of ψ on the circle C in the ζ -plane, where the potential $F(\zeta, t)$ has to be determined.

We notice first that it is only on C that the values are required, and as on C , $\zeta\bar{\zeta}=1$, the function of ζ that will reduce on C to the conjugate of z is easily seen to be $\bar{G}(1/\zeta)$. This function is holomorphic on and inside C , where it has the Laurent's expansion

$$\bar{z} = \bar{G} = \sum_{n=-1}^{\infty} \bar{A}_n \zeta^n; \quad (3.12)$$

we get also

$$\frac{d\bar{z}}{d\zeta} = \sum_{n=-1}^{\infty} \bar{A}_n \zeta^{n-1}. \quad (3.13)$$

Also putting in (3.11) $\zeta = e^{i\theta}$, we get the values of x and y on the circle C by separating real and imaginary parts, and then, inserting in (3.2) we get, on C :

$$\begin{aligned} \psi = & \frac{1}{2}U \sum_{n=-1}^{\infty} [i(\bar{A}_n - A_n) \cos n\theta + (\bar{A}_n + A_n) \sin n\theta] \\ & - \frac{1}{2}V \sum_{n=-1}^{\infty} [(\bar{A}_n + A_n) \cos n\theta + i(\bar{A}_n - A_n) \sin n\theta] \\ & - \frac{1}{2}\omega \left\{ \sigma_0 + \sum_{n=1}^{\infty} [(\bar{\sigma}_n + \sigma_n) \cos n\theta + i(\bar{\sigma}_n - \sigma_n) \sin n\theta] \right\}, \quad (3.21) \end{aligned}$$

where

$$\sigma_n = \sum_{m=-1}^{\infty} \bar{A}_m A_{m+n}.$$

Suppose that the complex potential $F(\zeta, t)$ is given by a Laurent's expansion

$$F(\zeta, t) = \sum_{n=1}^{\infty} B_n \zeta^{-n}, \quad (3.3)$$

convergent on and outside C .

The imaginary part of F must reduce on C to (3.21). Putting therefore in (3.3) $\zeta = e^{i\theta}$ and equating separately coefficients of $\sin n\theta$ and $\cos n\theta$ in the imaginary part of F to the corresponding coefficients in (3.21), the B 's are easily determined. We find

$$B_1 = WA_1 - \bar{W}\bar{A}_{-1} - i\omega\sigma_1,$$

$$B_n = WA_n - i\omega\sigma_n, \quad n=2, 3, \dots$$

In order to take account of a fluid circulation of strength Γ round the cylinder, we add to F a term

$$-\frac{i\Gamma}{2\pi} \log \zeta. \quad \dots \quad (3.31)$$

A function that will reduce on the circle C to the conjugate of F , will be

$$\bar{F} = \sum_{n=1}^{\infty} \bar{B}_n \zeta^n - \frac{i\Gamma}{2\pi} \log \zeta; \quad \dots \quad (3.4)$$

the series here is convergent on and inside C . Also

$$F + \bar{F} = \sum_{n=1}^{\infty} [B_n \zeta^{-n} + \bar{B}_n \zeta^n] - \frac{i\Gamma}{\pi} \log \zeta. \quad \dots \quad (3.5)$$

This series is convergent on C . In fact, it represents a function which is holomorphic in the ring space limited from inside by the circle of radius equal to the modulus of the singularity of $G(\zeta)$ nearest to C and from outside by the circle of radius equal to the inverse of that modulus with respect to C . This I call the ring C .

We get also

$$\begin{aligned} \frac{\partial F}{\partial \zeta} &= - \sum_{n=1}^{\infty} n B_n \zeta^{-n-1} - \frac{i\Gamma}{2\pi} \zeta^{-1}, \\ \frac{\partial \bar{F}}{\partial \zeta} &= \sum_{n=1}^{\infty} n \bar{B}_n \zeta^{n-1} - \frac{i\Gamma}{2\pi} \zeta^{-1}, \\ \frac{\partial F}{\partial t} &= \sum_{n=1}^{\infty} \dot{B}_n \zeta^{-n}, \quad \text{and} \quad \frac{\partial \bar{F}}{\partial t} = \sum_{n=1}^{\infty} \dot{\bar{B}}_n \zeta^n. \end{aligned}$$

We drop the term containing $\dot{\Gamma}$, otherwise the pressure would be multi-form owing to the presence of the term $\partial \phi / \partial t$ in the equation of the pressure.

4. Resultant Force and Resultant Couple.

These are best calculated in the ζ -plane, where equations (2.3) and (2.4) are written

$$Y + iX = -\frac{\rho}{2} \int_C \left\{ \left(\frac{\partial F}{\partial t} + \frac{\partial \bar{F}}{\partial t} \right) \frac{d\bar{z}}{d\zeta} + \left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{d\bar{z}} - (W - i\omega \bar{z}) \left(\frac{\partial F}{\partial \zeta} + \frac{\partial \bar{F}}{\partial \zeta} \right) \right\} d\zeta, \quad \dots \quad (4.1)$$

and

$$M = -\frac{\rho}{2} R \int_C z \left\{ \left(\frac{\partial F}{\partial t} + \frac{\partial \bar{F}}{\partial t} \right) \frac{d\bar{z}}{d\zeta} + \left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{d\bar{z}} - W \left(\frac{\partial F}{\partial \zeta} + \frac{\partial \bar{F}}{\partial \zeta} \right) \right\} d\zeta, \quad (4.2)$$

where integration is taken round the circumference C .

Supposing that the contour c does not contain singular points—in which case $\partial\zeta/dz$ is regular on C and consequently in the ring C —and inserting in (4.1) the appropriate values of the functions involved in the various integrals, we get

$$I_1 = \int_c \left(\frac{\partial \mathbf{F}}{\partial t} + \frac{\partial \bar{\mathbf{F}}}{\partial t} \right) \frac{d\bar{z}}{d\zeta} d\zeta = -2\pi i \left[\bar{A}_{-1} \dot{B}_1 - \sum_{n=1}^{\infty} n \bar{A}_n \dot{B}_n \right], \quad (4.11)$$

$$I_2 = \int_c \left(\frac{\partial \mathbf{F}}{\partial \zeta} \right)^2 \frac{d\zeta}{d\bar{z}} d\bar{z} = 0, \quad (4.12)$$

$$I_3 = -2[W - i\omega \bar{A}_0] \Gamma - 2\pi\omega \left[\bar{A}_{-1} \bar{B}_1 - \sum_{n=1}^{\infty} n \bar{A}_n \bar{B}_n \right]. \quad (4.13)$$

Collecting results and rearranging, we get

$$\begin{aligned} Y + iX = & \rho \Gamma [W - i\omega \bar{A}_0] + \rho \pi i \left[\bar{A}_{-1} \dot{B}_1 - \sum_{n=1}^{\infty} n \bar{A}_n \dot{B}_n \right] \\ & + \rho \pi \omega \left[\bar{A}_{-1} \bar{B}_1 - \sum_{n=1}^{\infty} n \bar{A}_n \bar{B}_n \right]. \quad (4.3) \end{aligned}$$

Or, rendering explicit the characteristics of the motion of the solid

$$\begin{aligned} Y + iX = & \rho \Gamma [W - i\omega \bar{A}_0] + \rho \pi i [2\dot{\bar{W}} \bar{A}_{-1} \bar{A}_1 - \dot{W} \{ \bar{A}_{-1} \bar{A}_{-1} + \sum n \bar{A}_n \bar{A}_n \} \\ & + i\dot{\omega} \{ \bar{A}_{-1} \bar{\sigma}_1 + \sum n \bar{A}_n \sigma_n \}] + \rho \pi \omega [2\bar{W} \bar{A}_{-1} \bar{A}_1 \\ & - W \{ \bar{A}_{-1} \bar{A}_{-1} + \sum n \bar{A}_n \bar{A}_n \} + i\omega \{ \bar{A}_{-1} \bar{\sigma}_1 + \sum n \bar{A}_n \sigma_n \}]. \quad (4.31) \end{aligned}$$

Also calling S the area of the cross-section of the cylinder and z_0 the affix of its centre of gravity, we have

$$S = \frac{1}{2} I \int \bar{z} dz = \pi \left[\bar{A}_{-1} \bar{A}_{-1} - \sum_{n=1}^{\infty} n \bar{A}_n \bar{A}_n \right],$$

and

$$Sz_0 = \frac{i}{2} \int \bar{z} z d\bar{z} = \pi \left[\bar{A}_{-1} \bar{\sigma}_1 - \sum_{n=1}^{\infty} n \bar{A}_n \sigma_n \right].$$

We therefore get

$$\begin{aligned} Y + iX = & \rho \Gamma W_0 + \rho [S - 2\pi \bar{A}_{-1} \bar{A}_{-1}] \{ i\dot{W} + \omega W \} + \rho [2\pi \bar{A}_{-1} \bar{\sigma}_1 - S \bar{z}_0] \\ & \times [i\omega^2 - \dot{\omega}] + 2\pi \rho \bar{A}_{-1} \bar{A}_1 [i\dot{\bar{W}} + \omega \bar{W}], \quad (4.32) \end{aligned}$$

where $W_0 = W - i\omega \bar{A}_0$ is the velocity of the point whose affix is A_0 , the hydrodynamic centre of the contour c . The first term represents the generalization of the theorem of Kutta-Joukowski to the varied motion of the cylinder, while the remaining terms represent inertia forces due to the acceleration of the cylinder.

In the same way the resultant moment is given by

$$\begin{aligned} M = R \left[\rho W C A_0 - \rho \pi i W \left\{ A_{-1} B_1 - \sum_{n=1}^{\infty} n A_n \bar{B}_n \right\} \right. \\ \left. - \rho \pi i \sum_{n=1}^{\infty} \left\{ \dot{B}_n \sum_{r=1}^{\infty} r \bar{A}_r A_{r-n} + \dot{B}_n \sum_{r=-1}^{\infty} r A_{r+n} \bar{A}_r \right\} \right]. \quad (4.4) \end{aligned}$$

Application.

Let us consider the elliptic profile

$$z = \lambda \zeta + \frac{\mu}{\bar{\zeta}},$$

where

$$\lambda = \frac{a+b}{2} \quad \text{and} \quad \mu = \frac{a-b}{2},$$

a and b being the major and the minor axes of the ellipse, we have

$$A_{-1} = \lambda, \quad A_1 = \mu,$$

and therefore,

$$\sigma_1 = 0, \quad \sigma_2 = \lambda\mu = (a^2 - b^2)/4,$$

$$B_1 = U(\mu - \lambda) - iV(\mu + \lambda) = -bU - iVa,$$

$$B_2 = -i\omega\lambda\mu = -i\omega(a^2 - b^2)/4.$$

Substituting in equations (4.3) and (4.4) we get, after reduction,

$$Y = \rho CU - \rho\pi a^2 \dot{V} - \rho\pi b^2 \omega U,$$

$$X = -\rho CV - \rho\pi b^2 \dot{U} + \rho\pi a^2 \omega V,$$

and

$$M = -\rho\pi(a^2 - b^2)UV - \frac{\pi}{8}\rho(a^2 - b^2)^2\dot{\omega},$$

the well-known results found by other methods.

5. *Angular Points, Cusps.*

Suppose that the contour c of the cross-section of the cylinder has an angular point of external angle $(\beta + 1)\pi$ at the point z_1 , where $|\beta| < 1$, and let ζ_1 be the affix of the corresponding point in the ζ -plane. In the vicinity of such a point $dz/d\zeta$ is of the form

$$\frac{dz}{d\zeta} = (\zeta - \zeta_1)^\beta \chi(\zeta), \quad \text{with } \chi(\zeta_1) \neq 0,$$

so that ζ_1 is a branch-point of this function and consequently of the functions z , \bar{z} , and $d/d\bar{z}\zeta$.

We see also that $\beta > 0$ corresponds to a corner pointing out into the liquid, while $\beta < 0$ corresponds to a re-entering corner.

As these singularities occur on the contour of integration they can be avoided by demi-circumferences of radius vanishingly small centred each at each singular point and placed either outside or inside C according as we are dealing with integrals involving the functions F or \bar{F} . Although the portions of the integrals I_1 , I_2 , I_3 and the similar ones in (4.2)

concerning these demi-circumferences have no contribution, the results (4.3) and (4.4) are not valid.

An exact knowledge of the singularities of z and $dz/d\zeta$ is necessary to perform these integrations.

When $\beta=1$, z_1 is a cusp, and consequently ζ_1 is a zero of $dz/d\zeta$ and an ordinary point of z , \bar{z} , and $d\bar{z}/d\zeta$.

ζ_1 is a pole of order one of the integrand in I_2 , so that, taking the principal value of this integral,

$$I_2 = -\pi i \operatorname{Res} \left[\left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right] \text{ at } \zeta = \zeta_1.$$

If the contour c contain n cusps at the points $z_j (j=1, 2, \dots, n)$, then we must add to (4.3) the terms

$$\frac{\rho}{2} \pi i \sum_{j=1}^n \operatorname{Res} \left[\left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right] \text{ at the poles } \zeta_j. \quad (4.5)$$

In the same way and for the same reasons we add to (4.4) the terms

$$\frac{\rho}{2} R \pi i \sum_1^n \operatorname{Res} \left[z \left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right] \text{ at the poles } \zeta_j; \quad (4.6)$$

Application.

Consider the case of a flat plate whose cross-section is a straight line of length $2l$. The transformation function is

$$z = \frac{l}{2} \left(\zeta + \frac{1}{\zeta} \right), \text{ and consequently, } \frac{d\zeta}{dz} = \frac{2\zeta^2}{l(\zeta^2 - 1)}.$$

We see that $A_{-1} = A_1 = l/2$, and therefore

$$\sigma_2 = \frac{l^2}{4}, \quad B_1 = -i l V \quad \text{and} \quad B_2 = -\frac{i \omega l^2}{4}.$$

Substituting in (4.31) and (4.4) we get, after reduction,

$$Y + iX = \rho \Gamma W - \rho \pi l^2 \dot{V} + \rho \pi i \omega l^2 V, \quad (4.7)$$

and,

$$M = -\rho \pi l^2 UV - \frac{\rho \pi l^4}{8} \frac{d\omega}{dt}. \quad (4.8)$$

The additive terms (4.5) become

$$\frac{\rho \pi i}{2} \sum \operatorname{Res} \left[\left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right] \text{ at } \zeta = 1 \text{ and at } \zeta = -1 = i \rho \Gamma V - i \rho \pi \omega l^2 V, \quad (4.71)$$

so that, adding these terms to (4.7), we get

$$\begin{aligned} X &= 0, \\ Y &= \rho \Gamma U - \rho \pi l^2 \dot{V}. \end{aligned} \quad (4.9)$$

The additive terms (4.6) are easily found to be identically zero.

Equations (4.9) show that the resultant pressure is always normal to the plate. It is a generalization of a result due to U. Cisotti (1927), in the case of liquid streaming steadily past, and circulating round a flat plate.

If the plate describes a circular path with constant speed $h\omega$ and constant angle of incidence α , so that

$$U = h\omega \cos \alpha, \quad V = -h\omega \sin \alpha,$$

then

$$X = 0, \quad Y = \rho \Gamma h \omega \cos \alpha,$$

and

$$M = \rho \pi l^2 h^2 \omega^2 \sin \alpha \cos \alpha.$$

5. Interpretation of the Additive Terms (4.5) and (4.6).

Suppose that the contour c has a cusp at z_1 and let $\zeta = e^{i\gamma}$ be the corresponding point on the circle C . The corresponding term in (4.5) represents a force (X_1, Y_1) (W. G. Bickley, 1929) such that

$$\begin{aligned} Y_1 + iX_1 &= \frac{\rho\pi i}{2} \operatorname{Res} \left[\left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right] \text{ at } \zeta = \zeta_1, \\ &= \frac{\rho\pi i}{2} \left[\left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} (\zeta - \zeta_1) \right]_{\zeta = \zeta_1} \dots \dots \dots (5.11) \end{aligned}$$

But in the vicinity of ζ_1 ,

$$\frac{dz}{d\zeta} = (\zeta - \zeta_1) \chi(\zeta), \quad \dots \dots \dots (5.12)$$

therefore

$$Y_1 + iX_1 = \frac{\rho\pi i}{2} \left(\frac{\partial F}{\partial \zeta_1} \right)^2 \frac{1}{\chi(\zeta_1)},$$

where $\frac{\partial F}{\partial \zeta_1}$ denotes the value of $\frac{\partial F}{\partial \zeta}$ at $\zeta = \zeta_1$, and

$$\chi(\zeta_1) = \chi(e^{i\gamma}) = \alpha_1 + i\beta_1.$$

Now, in the plane z the absolute velocity of the fluid with respect to the moving axes is

$$u - iv = \frac{\partial f}{\partial z} = \frac{\partial f'}{\partial \zeta} + (W - i\omega \bar{z}),$$

where f' is the complex relative potential.

Also in the ζ -plane we must have

$$\frac{\partial F}{\partial \zeta} = \frac{\partial F'}{\partial \zeta} + (W - i\omega \bar{z}) \frac{dz}{d\zeta},$$

where $\frac{d\mathbf{F}'}{d\zeta}$ represents the complex relative velocity. At a point on the circle C, we put: $\left| \frac{\partial \mathbf{F}}{\partial \zeta} \right| = q$, then,

$$\frac{d\mathbf{F}'}{d\zeta} = q e^{-i\left(\theta + \frac{\pi}{2}\right)} = -iq e^{-i\theta},$$

so that at the point $\zeta = \zeta_1$,

$$\frac{d\mathbf{F}}{d\zeta_1} = \frac{d\mathbf{F}'}{d\zeta_1} = -iq_1 e^{-i\gamma}, \quad \text{where } q_1 = \left| \frac{d\mathbf{F}}{d\zeta_1} \right|,$$

and consequently

$$Y_1 + iX_1 = -\frac{\rho\pi i}{2} q_1^2 e^{-2i\gamma} \frac{\alpha_1 - i\beta_1}{\alpha_1^2 + \beta_1^2}, \quad \dots \dots \dots (5.13)$$

and if Φ be the angle that this force makes with the x -axis, then

$$\tan \Phi = \frac{Y_1}{X_1} = \frac{\beta_1 \cos 2\gamma + \alpha_1 \sin 2\gamma}{\alpha_1 \cos 2\gamma - \beta_1 \sin 2\gamma}. \quad \dots \dots \dots (5.14)$$

The corresponding term in the expression of the resultant moment is

$$\begin{aligned} M_1 &= R \frac{\rho\pi i}{2} \left[\text{Res } z \left(\frac{\partial \mathbf{F}}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right] \text{ at } \zeta = \zeta_1, \\ &= R \frac{\rho\pi i}{2} z_1 \left(\frac{\partial \mathbf{F}}{\partial \zeta_1} \right)^2 \frac{1}{\chi(\zeta_1)}, \\ &= R(x_1 + iy_1)(Y_1 + iX_1) = x_1 Y_1 - y_1 X_1, \quad \dots \dots \dots (5.2) \end{aligned}$$

which is the moment about the origin of a force (X_1, Y_1) passing by the point z_1 of coordinates (x_1, y_1) . This force makes with the x -axis an angle Φ given by (5.14).

On the other hand, in the vicinity of the point ζ_1 , on the circle C,

$$\frac{dz}{d\zeta} = \frac{dz}{d\theta} \times \frac{d\theta}{d\zeta} = (e^{i\theta} - e^{i\gamma})\chi(e^{i\theta}),$$

so that

$$\frac{dz}{d\theta} = \frac{dx}{d\theta} + i \frac{dy}{d\theta} = ie^{i\theta} (e^{i\theta} - e^{i\gamma})(\alpha + i\beta). \quad \dots \dots \dots (5.3)$$

Also the tangent to c at a point near z_1 makes with the x -axis an angle ϕ defined by

$$\begin{aligned} \tan \phi &= \frac{dy}{dx} = \frac{dy}{d\theta} \div \frac{dx}{d\theta} \\ &= \frac{\beta[\sin 2\theta - \sin(\theta + \gamma)] - \alpha[\cos 2\theta - \cos(\theta + \gamma)]}{\beta[\cos 2\theta - \cos(\theta + \gamma)] + \alpha[\sin 2\theta - \sin(\theta + \gamma)]}. \end{aligned}$$

At the cusp, $\theta = \gamma$, so that $tg\phi$ appears to be indeterminate. But by Hospital's theorem we find

$$tg\phi = \frac{\alpha_1 \sin 2\gamma + \beta_1 \cos 2\gamma}{\alpha_1 \cos 2\gamma - \beta_1 \sin \gamma},$$

so that

$$\phi = \Phi \pm k\pi.$$

We see thus that the line of action of the force (X_1, Y_1) is the tangent to the contour c at the cusp. This force is due to the infinite pressure, because of the infinite velocity, produced at the cusp.

Consequently, the cusped profiles again do not follow the generalized theorem of Kutta-Joukowski. This theorem needs modification; I propose the following:—

In addition to inertia forces due to the acceleration of the profile and to the Kutta-Joukowski's lifting force, the cusped profiles are subject to a system of forces; along the tangent to the profile at each cusp a finite force (X_j, Y_j) acts and is such that

$$Y_j + iX_j = \frac{\rho\pi i}{2} \text{Res} \left[\left(\frac{\partial F}{\partial \zeta} \right)^2 \frac{d\zeta}{dz} \right]$$

at the point ζ_j corresponding to the cusp.

Summary.

In this paper the Blasius contour integral expressions for the determination of the resultant force and moment on a profile in a Joukowski's stream are extended to the case of the general uniplanar motion of the profile.

The potential function of the fluid motion induced by the motion of the solid is given in the form of a Laurent's series, the coefficients of which are expressed in terms of the coefficients of the expansion of the function transforming conformally the exterior of the profile into the exterior of a unit circle.

The resultant force and moment are calculated in terms of these coefficients, and the application to the classical case of the ellipse gives well-known results.

The consideration of cusped profiles leads to a modification of the generalized Kutta-Joukowski theorem—a modification which is dealt with.

Application to the case of a flat plate is given.

References.

- A. Masotti, 'Atti dei Lincei, Rendiconti,' iii. p. 472 (1926); v. p. 872 (1927).
 E. Carafoli, 'Aérodynamique des Ailes d'Avions,' Paris (1928).
 J. Pérés, 'Cours de Mécanique des Fluides' (1936).

- Rosa M. Morris, *Proc. Camb. Phil. Soc.* xxxiii. part 4 (1937) ; *Proc. Roy. Soc. London, A*, dccccvi. pp. 406-419 (1937) ; *Phil. Mag.* xxiii. and xxiv. (1937).
- M. A. Omara, *Proc. International Congress Math. Oslo* (1936) ; Report, Faculty of Science, Cairo (1934-35).
- U. Cisotti, ' *Atti dei Lincei, Rendiconti*,' v. p. 16 (1927).
- H. Villat, ' *Mécanique des Fluides* ' (1930).
- W. G. Bickley, *Phil. Trans. A*, ccxxviii. p. 235 (1929).
- H. Lamb, ' *Hydrodynamics* ' (1932).

The equation of continuity

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

gives $f' + 4r\phi + r^2\phi' = 0, \quad . \quad . \quad . \quad . \quad . \quad (3)$

dashes denoting differentiation with respect to r .

The equations of motion for a liquid are

$$\frac{D}{Dt}(u, v, w) = (X, Y, Z) - \frac{1}{\rho} \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) p + \nu \nabla^2(u, v, w), \quad . \quad (4)$$

On using (1) we find

$$\frac{Du}{Dt} = x \left[\frac{ff''}{r} + 2f\phi + rf'\phi + x^2 \left(\frac{f\phi'}{r} + 2\phi^2 + r\phi\phi' \right) \right], \quad . \quad . \quad (5.1)$$

$$\frac{Dv}{Dt} = y \left[f\phi + x^2 \left(\frac{f\phi'}{r} + 2\phi^2 + r\phi\phi' \right) \right], \quad . \quad . \quad . \quad . \quad (5.2)$$

$$\frac{Dw}{Dt} = z \left[f\phi + z^2 \left(\frac{f\phi'}{r} + 2\phi^2 + r\phi\phi' \right) \right], \quad . \quad . \quad . \quad . \quad (5.3)$$

and $\nabla^2 u = \frac{\partial}{\partial x} \left[x \left(\frac{f'}{r} + 9\phi + 2r\phi' \right) \right] + \frac{2f'}{r} + f'' - 3\phi - r\phi', \quad . \quad (5.4)$

$$\nabla^2 v = \frac{\partial}{\partial y} \left[y \left(\frac{f'}{r} + 9\phi + 2r\phi' \right) \right], \quad . \quad . \quad . \quad . \quad (5.5)$$

$$\nabla^2 w = \frac{\partial}{\partial z} \left[z \left(\frac{f'}{r} + 9\phi + 2r\phi' \right) \right]. \quad . \quad . \quad . \quad . \quad (5.6)$$

Hence, if we put

$$\lambda = -\frac{p}{\rho} + \nu x \left(\frac{f'}{r} + 9\phi + 2r\phi' \right) - \int rf\phi \, dr - x^2 \int (f\phi' + 2r\phi^2 + r^2\phi\phi') \, dr, \quad (6)$$

the equations of motion become

$$\frac{\partial \lambda}{\partial x} + \nu \left(\frac{2f'}{r} + f'' - 3\phi - r\phi' \right) - x \left(\frac{ff''}{r} + f\phi + rf'\phi \right) + 2x \int (f\phi' + 2r\phi^2 + r^2\phi\phi') \, dr + X = 0, \quad . \quad (7.1)$$

$$\frac{\partial \lambda}{\partial y} + Y = 0, \quad . \quad . \quad . \quad . \quad (7.2)$$

$$\frac{\partial \lambda}{\partial z} + Z = 0, \quad . \quad . \quad . \quad . \quad (7.3)$$

These are all satisfied if

$$Y = Z = 0,$$

and

$$\begin{aligned} X &= -\nu \left(2\frac{f'}{r} + f'' - 3\phi - r\phi' \right) + x \left(\frac{ff''}{r} + f\phi + rf'\phi \right) - 2x \int (f\phi' + 2r\phi^2 + r^2\phi\phi') \, dr \\ &= -\frac{\nu}{r^2} \frac{d}{dr} [r^2(f' - r\phi)] + x \left(\frac{f}{r} + r\phi \right) (f' - r\phi) + 2x \int \phi(f' - r\phi) \, dr. \end{aligned} \quad (8)$$

X also vanishes if $f' = r\phi$, (9)

which, combined with (3), gives the irrotational solution for a sphere moving through a non-viscous liquid. If U be its velocity in the direction of the x -axis, we know that the corresponding values of u, v, w are

$$u = -\frac{1}{2}U\left(\frac{a}{r}\right)^3 + \frac{3}{2}U\frac{x^2}{a^2}\left(\frac{a}{r}\right)^5, v = \frac{3}{2}U\frac{xy}{a^2}\left(\frac{a}{r}\right)^5, w = \frac{3}{2}U\frac{y^2}{a^2}\left(\frac{a}{r}\right)^5. \quad (9.1)$$

As these values do not satisfy the conditions at the surface of the sphere we cannot use this solution for viscous motion. The above is only a particular case of the general proposition that each non-viscous irrotational solution of the equations of motion corresponds to a viscous solution, the difficulty in the latter case being the satisfying of the boundary conditions. In general, if $X=Y=Z=0$, we must have

$$\frac{1}{r^2} \frac{d}{dr} [r^2(f' - r\phi)] = A, \quad (10.1)$$

$$\left(\frac{f}{r} + r\phi\right)(f' - r\phi) + 2 \int \phi(f' - r\phi) dr = B, \quad (10.2)$$

A and B being constants. The corresponding solution satisfying (7) is

$$\lambda + Arx - \frac{1}{2}Bx^2 = \text{a constant}. \quad (11)$$

Solving (3) and (10.1) together we get

$$\phi = \frac{D}{r^5} - \frac{1}{2} \frac{C}{r^3} - \frac{1}{15} A, f = -\frac{1}{3} \frac{D}{r^3} - \frac{1}{2} \frac{C}{r} + \frac{2}{15} Ar^2 + E, \quad . . (12.1)$$

and (3) and (10.2) we have

$$(f + r^2\phi)(r\phi'' + 6\phi') = 0. \quad (11.2)$$

$$f + r^2\phi = 0 \text{ only gives } f = \phi = 0,$$

$$r\phi'' + 6\phi' = 0 \text{ gives } f = -\frac{1}{3} \frac{D_1}{r^3} + \frac{2}{15} A_1 r^2 + E_1, \phi = \frac{D_1}{r^5} - \frac{1}{15} A_1. \quad . . (12.3)$$

The only solution common to (3), (10.1), (10.2) is

$$f = -\frac{1}{3} \frac{D}{r^3} + \frac{2}{15} Ar^2 + E, \phi = \frac{D}{r^5} - \frac{1}{15} A. \quad (13)$$

These can be made to satisfy the conditions at the surface of the sphere, but not at infinity. If we put $D=0$ for inside the sphere, and $A=0$ for the outside, we get the case of Hill's spherical vortex.

If we put the first term in the value of X given by (8) equal to zero, we get

$$\frac{d}{dr} [r^2(f' - r\phi)] = 0,$$

which combined with (3) gives

$$u = \frac{3}{4} U \frac{a}{r} + \frac{1}{4} U \left(\frac{a}{r} \right)^3 - \frac{3}{4} U \frac{x^2}{a^2} \left[\left(\frac{a}{r} \right)^5 - \left(\frac{a}{r} \right)^3 \right], \quad (14.1)$$

$$v = -\frac{3}{4} U \frac{xy}{a^2} \left[\left(\frac{a}{r} \right)^5 - \left(\frac{a}{r} \right)^3 \right], \quad (14.2)$$

$$w = -\frac{3}{4} U \frac{xy}{a^2} \left[\left(\frac{a}{r} \right)^5 - \left(\frac{a}{r} \right)^3 \right]. \quad (14.3)$$

This is Stokes's solution. The corresponding value of X is

$$X = \frac{3}{8} \frac{U^2}{a} \cdot \frac{x}{a} \left[\left(\frac{a}{r} \right)^6 - \frac{9}{2} \left(\frac{a}{r} \right)^4 \right] \quad (15.1)$$

If the sphere is fixed the new value of X is given by

$$X_1 = \frac{3}{8} \frac{U^2}{a} \frac{x}{a} \left[\left(\frac{a}{r} \right)^6 - \frac{9}{2} \left(\frac{a}{r} \right)^4 + 4 \left(\frac{a}{r} \right)^3 \right]. \quad (15.2)$$

At a great distance from the sphere the viscous forces ($\mu \nabla^2 u$ etc.) are of order $(\mu U a / r^3)$, ρX_1 of $(\rho U^2 a / r^2)$, and ρX of $(\rho U^2 a^2 / r^2)$. The ratio in the first case is the Reynolds number (Ur/ν) , which increases with r . But in the second case it is (Ua/ν) , which does not increase with r . And, since (Ua/ν) is supposed to be small in Stokes's solution, it should hold good even at a distance from the sphere if it is moving uniformly through the liquid. Oseen's criticism is therefore true only when the liquid flows past the fixed sphere.

Again, (15.1) shows that X can be neglected throughout the liquid if U^2 is very small compared with a . In other words we can also use the solution for a *big* sphere moving very slowly through the liquid.

But Stokes's solution makes the kinetic energy of the field infinite. To get a finite value for the energy we put

$$\rho X = -A\mu \frac{(n-2)(n-4)}{r^{n+1}} + \rho x \psi(r), \quad (16)$$

where A is a constant, $\psi(r)$ a function of r only, and n a positive quantity not equal to 4. $n=2$ only gives the solution in (14).

Comparing (16) with (8), we get

$$f' - r\phi = -\frac{A(n-4)}{r^n}, \quad (17.1)$$

$$\phi = \frac{D}{r^5} - \frac{A}{r^{n+1}}, \quad (17.2)$$

$$f = -\frac{1}{3} \frac{D}{r^3} + \frac{A(n-3)}{(n-1)r^{n-1}}, \quad (17.3)$$

D being another constant.

Adjusting A and D so that $u=U$, $v=w=0$, over $r=a$, we have

$$u = -\frac{1}{2} \frac{n-1}{n-4} U \left(\frac{a}{r}\right)^3 + \frac{3}{2} \frac{n-3}{n-4} U \left(\frac{a}{r}\right)^{n-1} + \frac{3}{2} \frac{n-1}{n-4} U \frac{x^2}{a^2} \left[\left(\frac{a}{r}\right)^5 - \left(\frac{a}{r}\right)^{n+1} \right], \quad \dots \quad (18.1)$$

$$v = \frac{3}{2} \frac{n-1}{n-4} U \frac{xy}{a^2} \left[\left(\frac{a}{r}\right)^5 - \left(\frac{a}{r}\right)^{n+1} \right], \quad \dots \quad (18.2)$$

$$w = \frac{3}{2} \frac{n-1}{n-4} U \frac{xz}{a^2} \left[\left(\frac{a}{r}\right)^5 - \left(\frac{a}{r}\right)^{n+1} \right], \quad \dots \quad (18.3)$$

while X becomes

$$\rho X = -\frac{3}{2} \frac{\mu U}{a^2} (n-1)(n-2) \left(\frac{a}{r}\right)^{n+1} - \frac{3}{2} \rho \frac{U^2}{a} \frac{n-1}{n-4} \frac{x}{a} \left[\frac{n^2-1}{n+4} \left(\frac{a}{r}\right)^{n+4} - \frac{3n+1}{2n} \left(\frac{a}{r}\right)^{2n} \right]. \quad \dots \quad (19)$$

The value of the drag due to pressures of the liquid on the solid is found to be

$$D = 6\pi\mu U a(n-1), \quad \dots \quad (20)$$

and the vorticity is given by

$$\xi=0, \eta = -\frac{3}{2} (n-1) \frac{Uz}{a^2} \left(\frac{a}{r}\right)^{n+1}, \zeta = \frac{3}{2} (n-1) \frac{Uy}{a^2} \left(\frac{a}{r}\right)^{n+1}. \quad \dots \quad (21)$$

If v/a is of the same order of small quantities as U , and (U^2/a) can be neglected, we see from (19) that the external force X can be taken to be negligible at all points of the liquid provided n is kept as small as possible. Further, if the kinetic energy is not to be infinite, we see from (18) that n should be chosen greater than $5/2$. Thus it appears from (2) that the value of the drag for a big sphere moving slowly through a slightly viscous liquid should be at least greater than one and a half times that given by Stokes.

In such a case R , the Reynolds number, is finite, and neither very small nor very great. If Zahm's ⁽⁸⁾ empirical formula

$$D = \pi \rho a^2 U^2 [14R^{-.85} + 0.24]$$

is adopted, we get the table on p. 217 for R lying between 4 and 8. Let

$$k_D = D/\pi \rho a^2 U^2,$$

$$k_{D1} = \frac{3}{2} \times \text{Stokes's value}, k_{D1} = \text{Zahm's value}.$$

The inertia terms are now as important as those due to viscosity.

Let us see what happens to the solution in (18) when n becomes very great. a/r is always less than one except on the surface of the sphere.

u, v, w therefore vanish on the sphere, but at all other points they approach the values

$$u = -\frac{1}{2}U\left(\frac{a}{r}\right)^3 + \frac{3}{2}U\frac{x^2}{a^2}\left(\frac{a}{r}\right)^5, \quad v = \frac{3}{2}U\frac{xy}{a^2}\left(\frac{a}{r}\right)^5, \quad w = \frac{3}{2}\frac{Uxz}{a^2}\left(\frac{a}{r}\right)^5,$$

which is the irrotational solution given in (9.1). X also becomes very small except on the surface of the sphere, where it is given by

$$(\rho X)_{r=a} = -\frac{3\mu U}{2a^2}(n-1)(n-2) - \frac{3}{2}\rho\frac{U^2x}{a}\frac{(n^2-1)(2n^2-5n-12)}{2n(n^2-16)}. \quad (22)$$

If ν and U are of order $1/n$, and a of order n , we see that (22) is of order $1/n^2$. (18) show that u, v, w are of order $1/n$. The vorticity given in (21) vanishes at all points except on the surface of the sphere, where it is of order $1/n$. The kinetic energy approaches the same value as in the corresponding irrotational motion, and is of order n .

R.	k_{D_0} .	k_{D_1} .
4	4.50	4.55
5	3.60	3.81
6	3.00	3.29
8	2.25	2.63

The drag in (20) approaches a finite value. Thus we see that, if $1/n^2$ and higher powers of $1/n$ can be neglected, (18) can give the solution for a very big sphere moving slowly through a slightly viscous liquid.

In this case the motion is everywhere irrotational except in a small layer near the surface of the sphere. The Reynolds number now is very large, and is of order n .

From the above discussion it appears that as the value of the Reynolds number, R , increases from 0 to ∞ the kinetic energy of the fluid decreases from infinity to the value we get for the corresponding irrotational motion. It will be seen that these results hold good for a circular cylinder as well.

These results are interesting in view of the boundary layer theory, in which an irrotational flow is assumed for all the liquid except in a thin layer near the body.

(2) Motion of a Circular Cylinder.

The corresponding results for a cylinder can now easily be obtained. We take

$$u = f(r) + x^2 \phi(r), \quad . \quad . \quad . \quad . \quad . \quad (23.1)$$

$$v = xy \phi(r). \quad . \quad . \quad . \quad . \quad . \quad (23.2)$$

The equation of continuity gives

$$f' + 3r\phi + r^2\phi' = 0, \quad . \quad . \quad . \quad . \quad . \quad (24)$$

and the equations of motion become

$$\begin{aligned} \frac{\partial \lambda}{\partial x} + v \left(\frac{f'}{r} + f'' - 2\phi - r\phi' \right) - x \left(\frac{ff'}{r} + f\phi + rf'\phi \right) \\ + 2x \int (f\phi' + 2r\phi^2 + r^2\phi\phi') dr + X = 0, \quad . \quad . \quad . \quad (25.1) \end{aligned}$$

$$\frac{\partial \lambda}{\partial y} + Y = 0, \quad . \quad . \quad . \quad . \quad . \quad (25.2)$$

where

$$\lambda = -\frac{p}{\rho} + vx \left(\frac{f'}{r} + 7\phi + 2r\phi' \right) - \int rf\phi dr - x^2 \int (f\phi' + 2r\phi^2 + r^2\phi\phi') dr. \quad (26)$$

To satisfy (25) we should put

$$\rho X = -\frac{\mu}{r} \frac{d}{dr} [r(f' - r\phi)] + x \left(\frac{f}{r} + r\phi \right) (f' - r\phi) + 2x \int \phi(f' - r\phi) dr. \quad (27)$$

In this case, too, $f' = r\phi$ gives the corresponding irrotational solution, and

$$\frac{d}{dr} [r(f' - r\phi)] = 0$$

give the known solution ⁽⁹⁾

$$u = \frac{D}{a^2} \left[-\frac{1}{2} \frac{a^2}{r^2} + \log r - \frac{x^2}{r^2} \left(1 - \frac{a^2}{r^2} \right) \right], \quad . \quad . \quad . \quad (28.1)$$

$$v = -\frac{D}{a^2} \frac{xy}{r^2} \left(1 - \frac{a^2}{r^2} \right), \quad . \quad . \quad . \quad . \quad . \quad (28.2)$$

where

$$D = \frac{a^2 U}{\log a - \frac{1}{2}}$$

and

$$X = \frac{2U^2 x}{(\log a - \frac{1}{2})^2} \cdot \frac{\log r}{r^2} \quad . \quad . \quad . \quad . \quad . \quad (29)$$

The defect in the solution is that the velocity at a great distance from the sphere instead of vanishing becomes logarithmically infinite.

Putting $\rho X = -A\mu \frac{(n-1)(n-3)}{r^{n+1}} + \rho x \phi(r), \quad . \quad . \quad . \quad . \quad (30)$

where n is neither equal to one nor three, we find, as in the case of a sphere,

$$u = -U \frac{n-1}{n-3} \frac{a^2}{r^2} + 2U \frac{n-2}{n-3} \left(\frac{a}{r}\right)^{n-1} + 2U \frac{n-1}{n-3} \frac{x^2}{a^2} \left[\left(\frac{a}{r}\right)^4 - \left(\frac{a}{r}\right)^{n+1} \right], \quad (31.1)$$

$$v = 2U \frac{n-1}{n-3} \frac{xy}{a^2} \left[\left(\frac{a}{r}\right)^4 - \left(\frac{a}{r}\right)^{n+1} \right], \quad \quad (31.2)$$

and

$$\begin{aligned} \rho X = & -\frac{2\mu U}{a^2} (n-1) \left(\frac{a}{r}\right)^{n+1} - 2\rho \frac{U^2}{a} \frac{n-1}{n-3} \frac{x}{a} \left[\frac{(n-1)^2}{n+3} \left(\frac{a}{r}\right)^{n+3} \right. \\ & \left. - \frac{2}{n} \left(\frac{a}{r}\right)^{2n} \right] \quad (32) \end{aligned}$$

The drag is independent of a , and is given by

$$D = 4\mu\pi U(n-1). \quad \quad (33)$$

By making v/a and U both very small compared with a we can make X negligible. n , of course, should be chosen as low as possible. If the kinetic energy is to be finite we see from (31) that n should be greater than 2. Thus for a big cylinder moving slowly through a slightly viscous liquid the value of the drag appears to be greater than

$$4\mu\pi U. \quad \quad (34)$$

The value obtained by Lamb also does not vary very much with a ⁽¹⁰⁾.

By making n sufficiently large we see from (31) that when a very big cylinder moves slowly through a slightly viscous liquid the motion is everywhere irrotational except in a thin layer near its surface.

SUMMARY.

The assumption of a uniform and symmetric flow has given the following results:—

- (1) Oseen's criticism of Stokes's solution that it is defective at points distant from the sphere holds good only when the liquid flows past the fixed sphere, and not when the sphere is moving uniformly through it.
- (2) If the kinetic energy can be infinite it should be possible to use Stokes's solution when U^2/a is too small to be taken into account.
- (3) When a big sphere moves slowly through a slightly viscous liquid the value of the drag should be at least greater than one and a half times that given by Stokes. In the case of a circular cylinder it should be greater than $4\mu\pi U$. In these cases the Reynolds number, R , has a finite value.
- (4) As R increases from 0 to ∞ the kinetic energy of the fluid decreases from infinity to the value for the corresponding irrotational motion.
- (5) For a very big sphere or a cylinder moving slowly through a slightly viscous liquid the motion is everywhere irrotational except in a thin layer near the boundary. R is now very large.

References.

- (1) Oseen, *Arkiv. Mat. Astron. Fysik*, vi. no. 29 (1910).
- (2) Noether, *Zeit. Math. Physik*, lxii. pp. 1-38 (1913).
- (3) Lamb, *Phil. Mag.* (6) xxi. pp. 112-119 (1911).
- (4) Filon, *Proc. Roy. Soc. A*, cxiii. pp. 7-27 (1926) ; *Trans. Roy. Soc. A*, cccxxvii. pp. 93-135 (1928).
- (5) Bairstow and others, *Proc. Roy. Soc. A*, c. pp. 394-413 (1922) ; *Trans. Roy. Soc. A*, cccxxiii. pp. 383-432 (1923).
- (6) Goldstein, *Proc. Roy. Soc. A*, cxxiii. pp. 216-235 (1929).
- (7) Lamb, ' *Hydrodynamics*, ' 5th Edition, p. 579.
- (8) Zahm, *Report of Nat. Advisory Committee Aeronautics*, no. 247, p. 520 (1926).
- (9) Berry and Swain, *Proc. Roy. Soc. A*, cii. p. 766 (1923).
- (10) Lamb, ' *Hydrodynamics*, ' 5th Edition, p. 583.

Hindu College,
Delhi, India.

XXI. *A Geometric Derivation of the Second Order Wave Equation.*

By H. W. HASKEY, Ph.D.*

[Received August 4, 1938.]

I. *Introduction.*

THE Riemannian five-dimensional world has been used to describe simultaneously gravitational and electromagnetic fields, and the metric is given by

$$d\sigma^2 = \gamma_{\mu\nu} dx^\mu dx^\nu, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\mu, \nu = 1, 2, 3, 4, 5.$$

This corresponds to Einstein's $ds^2 = g_{mn} dx^m dx^n$, where $m, n = 1, 2, 3, 4$. On writing $\gamma_{55} = 1$ and $\gamma_{5m} = \beta\phi_m$, where

$$\beta = \frac{e}{m_0 c^2}$$

and ϕ_m is the electromagnetic potential, the equation of motion of an electron of mass m_0 carrying a charge e is the equation of motion of a null-geodesic in a *cylindrical* Riemannian five-world⁽¹⁾.

A further development has resulted from the suggestion that the five-metric is described more fundamentally by the matrix-tensors γ^μ than by the tensors $\gamma^{\mu\nu}$. The relation between these quantities is

$$2\gamma^{\mu\nu} \cdot 1 = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu,$$

where 1 is the unit matrix, and, in addition, we write $2S^{\mu\nu} = \gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu$. The γ^μ 's are generalizations of the matrices used by Dirac in his wave-equation, and Flint has used them to define the matrix-length of a vector A, which is given by $\gamma^\mu A_\mu \psi \equiv \hbar \psi$, ψ being the scalar-matrix which gauges the five-world and is a function of the coordinates chosen to describe the Riemannian five-world. The matrix-length was introduced to provide a geometric basis for the quantum theory⁽²⁾.

Flint⁽²⁾ was able to derive, and not merely formulate, Dirac's first order wave equation geometrically as follows.

When a vector A undergoes parallel displacement its components change by $\Delta A_\mu = \Gamma_{\mu\nu}^\lambda A_\lambda dx^\nu$, where $\Gamma_{\mu\nu}^\lambda$ is the Riemannian five-dimensional

* Communicated by Dr. H. T. Flint.

three-index symbol. The change of matrix-length on parallel displacement is given by

$$\begin{aligned} \Delta(l\psi) \equiv \Delta(\gamma^\mu A_\mu \psi) &= \left\{ \left(\frac{\partial \gamma^\lambda}{\partial x^\nu} + \gamma^\mu \Gamma_{\mu\nu}^\lambda \right) \psi + \gamma^\lambda \frac{\partial \psi}{\partial x^\nu} \right\} A_\lambda dx^\nu \\ &= (K_\nu^\lambda \psi + \gamma^\lambda \psi_{,\nu}) A_\lambda dx^\nu, \end{aligned}$$

where $\psi_{,\nu}$ denotes ordinary differentiation of ψ in contrast to the symbol ; which denotes Riemannian covariant differentiation.

In general the change of matrix-length does not vanish for all vector components A_λ and all displacements, dx^ν , but by postulating that K_ν^λ vanishes on contraction, i. e., if $\lambda = \nu$, the equation

$$(\gamma^\lambda_\lambda \psi)_{;\lambda} \equiv \gamma^\lambda_\lambda \psi_{;\lambda} + K^\lambda_\lambda \psi = 0 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

is obtained. The K^λ_ν , which would normally be written as $\gamma^\lambda_{;\nu}$; ν are not zero, for, if they were, these fifteen vanishing quantities would make the Riemannian curvature zero and so upset the unification of gravitation and electromagnetism secured by use of the five-dimensional world. On equating to zero the single sum K^λ_λ , (2) gives

$$\gamma^m \frac{\partial \psi}{\partial x^m} + \gamma^5 \frac{\partial \psi}{\partial x^5} = 0. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$m = 1, 2, 3, 4.$$

Flint takes

$$\frac{\partial \psi}{\partial x^5} = \frac{2\pi m_0 c i}{h} \cdot \psi, \quad . \quad . \quad . \quad . \quad . \quad (5)$$

i. e., x^5 is treated as a cyclic coordinate, and the γ^m as the Tetrode matrices α^m in four dimensions. The remaining matrix, γ^5 , according to the theory, has the value $\gamma^5 = \alpha_0 - \beta \phi_m \alpha^m$, and this, with (5), converts (4) into Dirac's equation. The matrix α_0 has its square equal to unity and it anti-commutes with α^m .

In the present paper the foregoing geometric conceptions will be used in a non-Riemannian world of five dimensions obeying a modified Weyl metric to show that the second order wave equation accounting for spin, viz.,

$$\begin{aligned} W\psi &= \square^2 \psi - \frac{4\pi e i}{hc} (\phi^l \psi_{,l} + \phi^l_{,l} \psi) - \frac{2\pi e i}{hc} s^{ml} F_{ml} \psi \\ &\quad - \frac{4\pi^2 e^2}{h^2 c^2} \phi_l \phi^l \psi - \frac{4\pi^2 m_0^2 c^2}{h^2} \psi = 0, \end{aligned} \quad (6)$$

where summation over l and m for 1, 2, 3, 4, is implied, can be expressed as

$$[\gamma^\mu (\gamma^\lambda \psi)_{;\mu}]_{;\lambda} = -8/5 (\gamma^\mu \gamma^\lambda P_{\mu\lambda} - P) \psi, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where μ and λ are summed over the values 1 to 5. The expression $\square^2\psi$ denotes the Laplacian in four dimensions. Further, it will be shown that Dirac's equation can be written in the form

$$(\gamma^\mu\psi)_{;\mu} = (\gamma^\mu\psi)_{;\mu} = 0. \quad (7a)$$

The suffix $_{;\mu}$ is used to denote covariant differentiation in the continuum with the modified metric corresponding to that denoted by $_{;\mu}$ in the Riemannian continuum.

In (7) $P_{\mu\lambda}$ is the analogue of the Riemann-Christoffel tensor and P denotes the scalar curvature $\gamma^{\mu\lambda}P_{\mu\lambda}$.

The expression $\gamma^\mu\gamma^\lambda P_{\mu\lambda}$ may be referred to as the matrix curvature, and, like similar expressions introduced into this theory⁽⁴⁾, occupies a position corresponding to that of curvature in the geometry of large-scale phenomena. Eddington makes use of a similar expression in his 'Relativity Theory of Protons and Electrons'⁽⁵⁾. In our view the wave equation is a modified curvature equation. This view has been held by Flint⁽⁴⁾ and Schroedinger⁽⁶⁾, but in the latter's treatment arbitrary operators and other quantities are introduced which, to use a well-known metaphor, make the quantum equations disguised intruders on the geometric stage rather than structural features of it.

The present derivation, founded on the study of the changes of matrix-length caused by parallel displacement, is geometrical. This is evident from the fact that (7a) is of the form of a vanishing divergence, while (7) is a generalization of $\text{div}(\text{grad } \psi) = \text{constant matrix } \psi$.

II. The Application to Atomic Physics.

(a) Dirac's Equation.

For the sake of symmetry in the cylindrical Riemannian five-world we write $1 = \gamma_{55} = \beta\phi_5$. This gives

$$\phi^5 = \frac{1}{\beta}, \quad (\gamma^5)^2 = \gamma^{55} = \beta^2\phi_m\phi^m + 1 = \beta^2\phi_\mu\phi^\mu.$$

The constant ϕ^5 and the ϕ^m are assumed independent of x^5 , and as $\gamma = |\gamma_{\mu\nu}| = g$, we have

$$\phi^\mu_{;\mu} = \frac{\partial(\phi^\mu \log \sqrt{g})}{\partial x^\mu} = \frac{\partial(\phi^m \log \sqrt{g})}{\partial x^m} + \frac{\partial(\phi^5 \log \sqrt{g})}{\partial x^5} = \frac{\partial(\phi^m \log \sqrt{g})}{\partial x^m}.$$

We first note that Dirac's equation can be derived by Flint's method in any non-Riemannian five-world for which

$$\Delta A_\mu = (F_{\mu\nu}^\lambda + T_{\mu\nu}^\lambda)A_\lambda dx^\nu,$$

where $\gamma^\lambda_{;\mu} \equiv \gamma^\lambda_{;\mu} + \gamma^\sigma T_{\sigma\mu}^\lambda$ by definition, provided that

$$\gamma^\lambda_{;\lambda} = 0. \quad (8)$$

This equation is the analogue of Flint's equation $K_{\lambda}^{\lambda} = \gamma^{\lambda}_{;\lambda} = 0$, and Dirac's equation assumes the form $(\gamma^{\lambda}\psi)_{;\lambda} = 0$.

We shall, however, select the arbitrary tensor $T_{\mu\nu}^{\lambda}$ to agree with the system of gauging suggested by the quantum theory⁽³⁾, and this requires

$$\Delta l^2 = -4bl^2\phi_{\mu}dx^{\mu}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where

$$b = \frac{\pi e i}{\hbar c}.$$

Since $2\gamma^{\mu\nu} = \gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu}$ and $\gamma^{\mu\nu}_{;\rho} = 0$, we find that

$$\gamma^{\nu}_{;\rho}\gamma^{\mu} + \gamma^{\mu}\gamma^{\nu}_{;\rho} + \gamma^{\mu}_{;\rho}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu}_{;\rho} = 2(\gamma^{\mu\sigma}T_{\sigma\rho}^{\nu} + \gamma^{\sigma\nu}T_{\sigma\rho}^{\mu}) \quad . \quad . \quad (10)$$

and that (9) holds if each member of (10) is equivalent to

$$-8b\gamma^{\mu\nu}\phi_{\rho}.$$

This can be satisfied by taking

$$T_{\rho\lambda}^{\mu} = b(-2\delta_{\rho}^{\mu}\phi_{\lambda} + n\delta_{\lambda}^{\mu}\phi_{\rho} - n\gamma_{\rho\lambda}\phi^{\mu}), \quad . \quad . \quad . \quad . \quad (12)$$

where n is any arbitrary number. From (8) and (12)

$$\gamma^{\lambda}_{;\lambda} = 2(1-2n)b\phi_{\mu}\gamma^{\mu}, \quad . \quad . \quad . \quad . \quad (13)$$

an equation which is really the foundation of Dirac's equation derived by a consideration of matrix-length changes of a vector in this continuum.

(b) *The Curvature.*

Writing $R_{\mu\nu}$ as the contracted Riemann-Christoffel tensor in a Riemannian five-world and using (13) and

$$P_{\mu\nu} = R_{\mu\nu} - T_{\mu\nu}^{\sigma}{}_{;\sigma} + T_{\mu\sigma}^{\sigma}{}_{;\nu} + T_{\rho\nu}^{\sigma}T_{\mu\sigma}^{\rho} - T_{\rho\sigma}^{\sigma}T_{\mu\nu}^{\rho}, \quad . \quad . \quad . \quad (14)$$

we find, after some calculation,

$$P_{\mu\nu} - R_{\mu\nu} = -2bF_{\mu\nu} + 3nb\phi_{\mu;\nu} + nb\gamma_{\mu\nu}\phi^{\sigma}_{;\sigma} + 3n^2b^2(\gamma_{\mu\nu}\phi^{\sigma}\phi_{\sigma} - \phi_{\mu}\phi_{\nu}),$$

where

$$F_{\mu\nu} \equiv \phi_{\mu;\nu} - \phi_{\nu;\mu}, \quad . \quad . \quad . \quad . \quad (15)$$

and

$$P - R = 8nb\phi^{\sigma}_{;\sigma} + 12n^2b^2\phi^{\sigma}\phi_{\sigma}, \quad . \quad . \quad . \quad . \quad (16)$$

while

$$\gamma^{\mu}\gamma^{\nu}P_{\mu\nu} - R = \frac{(3n-4)}{2}bs^{\mu\nu}F_{\mu\nu} + 8nb\phi^{\sigma}_{;\sigma} + 12n^2b^2\phi^{\sigma}\phi_{\sigma}, \quad . \quad . \quad (17)$$

and therefore

$$\gamma^{\mu}\gamma^{\nu}P_{\mu\nu} - P = \frac{(3n-4)}{2}bs^{\mu\nu}F_{\mu\nu}. \quad . \quad . \quad . \quad . \quad (18)$$

Equation (18) shows that the spin contribution $2bs^{\mu\nu}F_{\mu\nu}$ is merely proportional to the difference between the generalized and ordinary scalar curvature of the continuum.

(c) *The Generalized Divergence.*

To complete our proof that (7) is Schrodinger's equation we must evaluate the left-hand side of (7) and choose the arbitrary number n in (12) and (13) suitably. Now

$$[\gamma^\mu(\gamma^\lambda\psi); \mu; \lambda] = [\gamma^\mu\gamma^\lambda; \mu\psi + \gamma^\mu\gamma^\lambda\psi; \mu; \lambda]_{; \lambda} \\ = \gamma^\mu; \lambda\gamma^\lambda; \mu\psi + \gamma^\mu\gamma^\lambda; \mu; \lambda\psi + \gamma^\mu\gamma^\lambda; \mu\psi; \lambda + S^{\mu\lambda}; \lambda\psi_\mu + \bigtriangleup\psi. \quad (19)$$

where

$$\bigtriangleup\psi \text{ is } \gamma^{\mu\lambda}\psi; \mu; \lambda.$$

We now evaluate the terms in (19) using (13), the Riemannian $B_{\mu\nu\sigma\epsilon}$, and not making use of Dirac's first order equation $(\gamma^\mu\psi)_{; \mu} = 0$. If $B_{\mu\nu\sigma\epsilon}$ is the Riemann-Christoffel tensor (in the Riemannian five-world)

$$\gamma^\lambda; \mu; \lambda - \gamma^\lambda; \lambda; \mu = \gamma^{\rho\lambda}\gamma^\epsilon B_{\rho\mu\lambda\epsilon}, \quad \dots \quad (20) \\ \therefore \gamma^\mu\gamma^\lambda; \mu; \lambda - \gamma^\mu\gamma^\lambda; \lambda; \mu = \gamma^{\rho\lambda}\gamma^\mu\gamma^\epsilon B_{\rho\mu\lambda\epsilon} \\ = -\gamma^\mu\gamma^\epsilon\gamma^{\rho\lambda} B_{\epsilon\mu\lambda\rho}, \\ = -\gamma^\mu\gamma^\epsilon R_{\epsilon\mu},$$

$$\therefore \gamma^\mu\gamma^\lambda; \mu; \lambda - \gamma^\mu\gamma^\lambda; \lambda; \mu = -R, \quad \dots \quad (20 a)$$

$$\therefore \gamma^\mu\gamma^\lambda; \mu; \lambda = -R + 2(1-2n)b\gamma^\mu(\phi_\lambda\gamma^\lambda); \mu \text{ by (13)} \\ = -R + 2(1-2n)b[\gamma^\mu\phi_\lambda; \mu\gamma^\lambda + \gamma^\mu\phi^\lambda\gamma^\lambda; \mu].$$

Now

$$\gamma^\mu\gamma^\lambda = \gamma^{\mu\lambda}1 + s^{\mu\lambda}, \\ \therefore (\gamma^\mu\gamma^\lambda); \mu = 0 + s^{\mu\lambda}; \mu, \\ \therefore \gamma^\mu; \mu\gamma^\lambda + \gamma^\mu\gamma^\lambda; \mu = s^{\mu\lambda}; \mu.$$

Hence by (13),

$$\gamma^\mu\gamma^\lambda; \mu = s^{\mu\lambda}; \mu - 2(1-2n)b\phi_\mu\gamma^\mu\gamma^\lambda, \quad \dots \quad (21)$$

and, after a little reduction,

$$\gamma^\mu\gamma^\lambda; \mu; \lambda = 2(1-2n)b\phi^\mu; \mu + (2n-1)bs^{\mu\lambda}F_{\mu\lambda} \\ - R + 2(1-2n)b\phi_\lambda s^{\mu\lambda}; \mu + 4b^2(1-2n)^2\phi_\lambda\phi^\lambda. \quad \dots \quad (22)$$

Post-multiplying (20) by γ^μ we obtain, analogously to (20 a),

$$\gamma^\lambda; \mu; \lambda\gamma^\mu = -R + \gamma^\lambda; \lambda; \mu\gamma^\mu. \quad \dots \quad (20 b)$$

We can now show that $s^{\mu\lambda}; \lambda; \mu = 0$ without using (13). We have

$$s^{\nu\sigma}; \nu; \sigma - s^{\nu\sigma}; \sigma; \nu = s^{\nu\epsilon}\gamma^{\sigma\mu}B_{\mu\nu\sigma\epsilon} + s^{\sigma\epsilon}\gamma^{\nu\mu}B_{\nu\mu\sigma\epsilon}.$$

Remembering that $B_{\mu\nu\sigma\epsilon}$ is antisymmetric in μ, ϵ and in σ, ν , and is symmetric for the double interchange of μ, ν and σ, ϵ , we have

$$\begin{aligned} s^{\nu\sigma}_{;\nu;\sigma} - s^{\nu\sigma}_{;\sigma;\nu} &= -s^{\nu\epsilon}\gamma^{\sigma\mu}B_{\nu\epsilon\mu\sigma} + s^{\epsilon\sigma}\gamma^{\tau\nu}B_{\sigma\epsilon\tau\nu} \\ &= -s^{\nu\epsilon}R_{\nu\epsilon} + s^{\epsilon\sigma}R_{\sigma\epsilon} \\ &= 0 + 0. \end{aligned}$$

But $s^{\nu\sigma}$ is antisymmetric, hence $s^{\nu\sigma}_{;\nu;\sigma} = -s^{\sigma\nu}_{;\nu;\sigma}$, so that

$$2s^{\nu\sigma}_{;\nu;\sigma} = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

We have

$$(\gamma^\mu\gamma^\lambda)_{;\lambda;\mu} = \gamma^\mu_{;\lambda;\mu}\gamma^\lambda + \gamma^\mu_{;\lambda}\gamma^\lambda_{;\mu} + \gamma^\mu_{;\lambda}\gamma^\lambda_{;\mu} + \gamma^\mu_{;\lambda}\gamma^\lambda_{;\mu} + \gamma^\mu_{;\lambda}\gamma^\lambda_{;\mu}$$

and

$$(\gamma^\mu\gamma^\lambda)_{;\lambda;\mu} = (\gamma^{\mu\lambda} + s^{\mu\lambda})_{;\lambda;\mu} = s^{\mu\lambda}_{;\lambda;\mu}. \quad . \quad . \quad . \quad . \quad (24)$$

Using (13), (20 b), (22), and (23) we find, after simplification, that

$$\gamma^\mu_{;\lambda}\gamma^\lambda_{;\mu} = 4(2n-1)b\phi^\lambda_{;\lambda} + (4n-2)^2b^2\phi_\lambda\phi^\lambda + R. \quad . \quad . \quad . \quad (25)$$

The value of the last term of (19) is well known⁽⁷⁾,

$$\square\psi = \square\psi - 4b(\phi^l\psi_{,l} + \phi^l_{,l}\psi) + 4b^2\phi_l\phi^l\psi - \frac{4\pi^2m_0^2c^2}{h^2}\psi,$$

where $\square\psi$ is the Laplacian of four dimensional relativity. In order to obtain this result it is necessary to make use of Kaluza's relations for the coefficients $\gamma_{\mu\nu}$ and $\gamma^{\mu\nu}$. We then obtain by (6)

$$\square\psi = W\psi + 2bs^{ml}F_{ml}\psi. \quad . \quad . \quad . \quad . \quad . \quad (26)$$

(d) *The Derivation of the Second Order Equation.*

By means of (19), (21), (22), (25), and (26), we obtain

$$\begin{aligned} \{\gamma^\mu(\gamma^\lambda\psi)_{;\mu}\}_{;\lambda} &= W\psi + 2bs^{ml}F_{ml}\psi + (2n-1)b \\ &\quad \times \{(2\phi^\lambda_{;\lambda} + s^{\mu\lambda}F_{\mu\lambda} - 2\phi_\lambda s^{\mu\lambda}_{;\mu})\psi + 2\phi_\mu\gamma^\lambda\gamma^\mu\psi_{;\lambda}\}. \quad (19 a) \end{aligned}$$

We now set the arbitrary number n in (13) equal to $1/2$, and thus (19 a) becomes

$$\{\gamma^\mu(\gamma^\lambda\psi)_{;\mu}\}_{;\lambda} = W\psi + 2bs^{ml}F_{ml}\psi, \quad . \quad . \quad . \quad . \quad (19 b)$$

while (18) becomes

$$(\gamma^\mu\gamma^\lambda P_{\mu\lambda} - P)\psi = -5/4bs^{\mu\lambda}F_{\mu\lambda}\psi. \quad . \quad . \quad . \quad . \quad (18 a)$$

Now

$$\begin{aligned} s^{\mu\lambda}F_{\mu\lambda} &= s^{ml}F_{ml} + 2s^{m5}F_{m5} \\ &= s^{ml}F_{ml} + 2s^{m5}\left(\frac{\partial\phi_m}{\partial x^5} - \frac{\partial\phi_5}{\partial x^m}\right) \\ &= s^{ml}F_{ml}, \end{aligned}$$

since ϕ_s is constant and ϕ_m does not contain x^5 , the Riemannian world of five dimensions being cylindrical.

From (18 *a*), (19 *b*), and (27) we see that

$$\{\gamma^\mu(\gamma^\lambda\psi)_{;\mu}\}_{;\lambda} = W\psi - 8/5(\gamma^\mu\gamma^\lambda P_{\mu\lambda} - P)\psi. \quad (28)$$

Thus the wave equation (6), viz., $W\psi=0$ is equivalent to

$$\{\gamma^\mu(\gamma^\lambda\psi)_{;\mu}\}_{;\lambda} = -8/5(\gamma^\mu\gamma^\lambda P_{\mu\lambda} - P)\psi. \quad (30)$$

In this derivation it was noted that the first order equation of the quantum theory was not used. Instead we made use of the relation (13). In the usual derivation of the second order equation Dirac's first order equation is employed and is responsible for the introduction of the spin term. The derivation of the second order equation and introduction of the spin term carried out here is based on the relation (13).

For the form of metric which has been suggested here and which is defined by (12) with $n=\frac{1}{2}$ we have

$$\gamma^\mu_{;\mu} \equiv \gamma^\mu_{;\mu} + T^\lambda_{\mu\lambda} \gamma^\mu = 0.$$

The natural extension of the law $\gamma^\mu_{;\mu}=0$ suggested by Flint is $\gamma^\mu_{;\mu}=0$, and we see that with the modified Weyl metric chosen the two laws become the same, and thus that we may describe the quantum law in this more general way.

We can derive another relation satisfied by these matrices and connecting them with the Riemannian curvature by means of (20 *a*), (20 *b*), and (25). For since $\gamma^\lambda_{;\lambda}=0$, we have

$$\gamma^\mu\gamma^\lambda_{;\mu;\lambda} = \gamma^\lambda_{;\mu;\mu}\gamma^\mu = -R$$

and

$$\gamma^\mu_{;\lambda}\gamma^\lambda_{;\mu} = R.$$

Moreover, from (21) we obtain

$$\gamma^\mu\gamma^\lambda_{;\mu} = g^{\mu\lambda}_{;\mu}.$$

These relations are illustrative of the fundamental character played by these matrices in the geometry and metrics of the continuum.

IV. CONCLUSION.

In this way we make the second order wave equation equivalent to one which is a gauging equation and the function ψ appears as a gauging function or scale-factor. This supports the suggestion that the quantum theory is to be regarded as the expression of the fact that a particular space metric belongs to the physical world. It would appear that the idea of Weyl and Eddington has been applied in the wrong place in physics—the metric is not to be traced in electromagnetic phenomena, but in the finer world-structure that the quantum theory describes.

This view accords well with the new character of Planck's quantum of action which the later form of the quantum theory imparts to it.

References.

- (1) Fisher, Proc. Roy. Soc. A, cxxiii. p. 489.
- (2) Flint, Proc. Roy. Soc. A, cl. p. 421.
- (3) Flint, Proc. Roy. Soc. A, cl. p. 421 ; Proc. Phys. Soc. xlviii. p. 433.
- (4) Flint, Proc. Roy. Soc. A, cxliv. p. 413 ; A, cxlv. p. 645.
- (5) Eddington, 'Relativity Theory of Protons and Electrons,' ch. x.
- (6) Schroedinger, *Sitz. Preuss. Akad.* xi. p. 105.
- (7) O. Klein, *Zeit. f. Phys.* xxxvi. p. 895 (1926).

XXII. *The Lift and Drag of a Rectangular Wing spanning
a Free Jet of Circular Section.*

By H. B. SQUIRE, Manchester University *.

[Received September 14, 1938.]

SUMMARY.

Investigations by Glauert⁽¹⁾ and Stüper^{(2), (3)} lead to discordant results for the effect of the jet boundary on the lift and drag of a rectangular wing spanning a free jet of circular section. This problem is solved by an independent method; the results obtained are in agreement with Glauert's results.

1. *Introduction.*

PREVIOUS papers by Glauert⁽¹⁾ and Stüper^{(2), (3)} have dealt with the problem of the lift and induced drag of a rectangular untwisted wing spanning a free jet of circular section; the results obtained by these two authors are in disagreement. Since approximations of uncertain magnitude are introduced in each case, it seemed desirable to solve the problem by a third method to determine which of the two results is correct.

The scanty experimental data⁽⁴⁾ indicate that the correction to the lift curve slope due to the effect of the jet boundary is greater than the values given either by Glauert or by Stüper, so that this report does not give a final answer to this problem. The disagreement between theory and experiment may be due either to the effect of the jet boundary, which is not sharply defined, as is assumed in the analysis, or to the effect of the induced flow in distorting the shape of the jet. A preliminary investigation of the latter effect has been made by Lotz and Riegels⁽⁵⁾, who have determined the wind-tunnel corrections for a jet of triangular section.

2. *Formulation of the Problem.*

The wing is assumed to be of constant chord c and to be symmetrically placed across a free jet of radius R , so that the aspect ratio of the wing is $2R/c$. Taking the origin at the centre of the wing and the x -axis along

* Communicated by the Author.

the wing-span, the induced velocity at any point of the wing due to the trailing vortices is

$$w_1(x) = \frac{1}{4\pi} \int_{-R}^R \frac{dK}{dx'} \cdot \frac{dx'}{x-x'},$$

where K is the circulation round any section of the wing.

The boundary condition for a free jet requires that the tangential velocity at the circular boundary far downstream shall be zero, and this condition can be satisfied by introducing image vortices of the same strength and sign at the points inverse to the positions of the wing trailing vortices. Thus a trailing vortex of strength dK at a distance x' from the centre of the jet will have an image of strength dK at the point R^2/x' , and this will produce an additional downwash at a point x on the wing of amount

$$\frac{1}{4\pi} \frac{dK}{x - R^2/x'},$$

where the numerical factor 4π occurs instead of the value 2π associated with infinite line vortices because the trailing vortex system extends downstream only. By integrating over the whole span we obtain for the induced velocity of the image system

$$w_2(x) = \frac{1}{4\pi} \int_{-R}^R \frac{dK}{dx'} \frac{dx'}{x - R^2/x'}.$$

Hence the total downwash at the point x is

$$w(x) = w_1 + w_2 = \frac{1}{4\pi} \int_{-R}^R \frac{dK}{dx'} \left[\frac{1}{x-x'} + \frac{1}{x-R^2/x'} \right] dx'. \quad (1)$$

The usual relation between circulation and downwash is

$$2K = a_0 c (V\alpha - w), \quad (2)$$

where α is the angle of incidence of the wing measured from the zero lift angle, V is the stream velocity, and $a_0 (\doteq 6)$ is the slope of the lift curve for two-dimensional flow.

Substituting from (1) for w , equation (2) becomes

$$\frac{1}{4\pi V} \int_{-R}^R \frac{dK}{dx'} \left[\frac{1}{x-x'} + \frac{1}{x-R^2/x'} \right] dx' + \frac{2K}{a_0 c V} = \alpha. \quad (3)$$

If we put

$$\left. \begin{aligned} y &= x/R, \quad y' = x'/R, \\ K &= 4\pi V R \alpha \Gamma, \end{aligned} \right\}, \quad (4)$$

equation (3) becomes

$$\int_{-1}^1 \frac{d\Gamma}{dy'} \left[\frac{1}{y-y'} + \frac{y'}{yy'-1} \right] dy' + \mu \Gamma = 1, \quad (5)$$

where

$$\mu = \frac{8\pi R}{a_0 c}.$$

3. Method of Solution.

The principal difficulty in solving (5) is in the choice of a suitable series expansion for Γ . Since the circulation K must tend to zero at the wing-tips it follows from (2) that w is finite and equal to $V\alpha$ just inside the jet-boundary. But the condition of zero tangential velocity along the circular boundary far downstream requires that w shall vanish at the wing-tips, and hence w has a discontinuity at $x = \pm R$. The problem is therefore analogous to the case of a discontinuity in the incidence of a wing at some point of its span; for the latter problem equation (2) still applies, and thus a discontinuity in α requires that w shall be discontinuous, since K must be continuous. The problem of a discontinuity in incidence has been considered by Betz and Petersohn⁽⁶⁾, who proved that dK/dx is logarithmically infinite at the discontinuity, and their result suggests the correct form of the series expansion for K or Γ .

To solve (5) it is therefore assumed that Γ can be represented by an expression of the form *

$$\Gamma = B[2 \log 2 - (1-y) \log (1-y) - (1+y) \log (1+y)] - \sum_{n=0}^{\infty} \frac{b_n(1-y^{2n+2})}{2n+2}. \quad (6)$$

This expression is symmetrical about the centre of the wing and vanishes at the wing-tips $y = \pm 1$. It is verified below that the form (6) for Γ enables the relation (2) between the circulation and the downwash to be satisfied near the wing-tips. (The form assumed for Γ in Glauert's investigation leads to w_2 , the downwash induced by the image vortices, being infinite at the wing-tips, while Stüper's form leads to w , the total downwash, being zero at the wing-tips.)

Substituting from (6) in (5) we obtain

$$\int_{-1}^1 \left[B \log \frac{1-y'}{1+y'} + \sum_{n=0}^{\infty} b_n y'^{2n+1} \right] \left[\frac{1}{y-y'} + \frac{y'}{yy'-1} \right] dy' + \mu \left[B\{2 \log 2 - (1-y) \log (1-y) - (1+y) \log (1+y)\} - \sum_{n=0}^{\infty} \frac{b_n(1-y^{2n+2})}{2n+2} \right] = 1. \quad (7)$$

It is shown in the appendix that

$$\int_{-1}^1 \log \frac{1-y'}{1+y'} \left[\frac{1}{y-y'} + \frac{y'}{yy'-1} \right] dy' = \frac{\pi^2}{2} + \frac{1-y^2}{2y^2} \log^2 \frac{1-y}{1+y}. \quad (8)$$

* All logarithms are to base e .

for $-1 < y < 1$. Further, if $J_n(y)$ and $K_n(y)$ are defined by the relations

$$J_n(y) = - \int_{-1}^1 \frac{y'^{2n+1}}{y-y'} dy',$$
$$K_n(y) = - \int_{-1}^1 \frac{y'^{2n+2}}{yy'-1} dy',$$

it can easily be shown that

$$J_n(y) = y^{2n+1} \log \frac{1-y}{1+y} + 2 \left[\frac{1}{2n+1} + \frac{y^2}{2n-1} + \dots + \frac{y^{2n}}{1} \right],$$
$$-K_n(y) = \frac{1}{y^{2n+3}} \log \frac{1-y}{1+y} + 2 \left[\frac{1}{(2n+1)y^2} + \frac{1}{(2n-1)y^4} + \dots + \frac{1}{y^{2n+2}} \right].$$

In addition, $J_n(1) + K_n(1) = 0$. The functions $J_n(y)$ and $K_n(y)$ are tabulated in Table I. for a number of values of y , for $n=0, 1, 2$.

TABLE I.
Values of the Functions J_n and K_n .

y	0	0.1	0.2	0.3	0.4	0.5	0.6
J_0	2.0	1.9799	1.9189	1.8143	1.6611	1.4507	1.1682
J_1	0.6667	0.6865	0.7434	0.8300	0.9324	1.0293	1.0872
J_2	0.40	0.4069	0.4298	0.4747	0.5492	0.6573	0.7914
K_0	0.6667	0.6707	0.6831	0.7052	0.7390	0.7889	0.8625
K_1	0.40	0.4029	0.4118	0.4277	0.4523	0.4889	0.5439
K_2	0.2857	0.2880	0.2950	0.3073	0.3267	0.3557	0.3997
y	0.7	0.8	0.9	0.95			
J_0	0.7859	0.2422	-0.6500	-1.4804			
J_1	1.0517	0.8217	0.1402	-0.6694			
J_2	0.9154	0.9259	0.5136	-0.2041			
K_0	0.9752	1.1664	1.5699	2.0569			
K_1	0.6296	0.7809	1.1151	1.5404			
K_2	0.4686	0.5951	0.8828	1.2636			

Substituting from (8) and (9), equation (7) becomes

$$B \left[\frac{\pi^2}{2} + \frac{1-y^2}{2y^2} \log^2 \left(\frac{1-y}{1+y} \right) \right] - \sum_{n=0}^{\infty} b_n (J_n + K_n) + \mu \left[B \{ 2 \log 2 \right.$$
$$\left. - (1-y) \log (1-y) - (1+y) \log (1+y) \} - \sum_{n=0}^{\infty} \frac{b_n (1-y^{2n+2})}{2n+2} \right] = 1.$$

. . . (10)

This equation can be satisfied near the wing-tips for $y = \pm (1 - \epsilon)$, where ϵ is a vanishingly small quantity, if $B = 2/\pi^2$, thus verifying that the form

(6) gives the correct distribution of circulation near the tips. With this value of B equation (10) becomes

$$\frac{1-y^2}{\pi^2 y^2} \log^2 \left(\frac{1-y}{1+y} \right) + \frac{2\mu}{\pi^2} [2 \log 2 - (1-y) \log (1-y) - (1+y) \log (1+y)] - \Sigma b_n \left[J_n + K_n + \frac{\mu(1-y^{2n+2})}{2n+2} \right] = 0. \quad (11)$$

Equation (11) was solved approximately for $\mu=5, 10, 15$, corresponding roughly to aspect ratios of 2.5, 5, and 7.5, by retaining only three coefficients b_0, b_1, b_2 , the others being put equal to zero; these three coefficients were determined by satisfying (11) at the points $y=0, 0.6, 0.9$. By symmetry the equation is satisfied at $y=-0.6, -0.9$, and it is also satisfied near the tips. The numerical values found for the coefficients are given in Table II.

TABLE II.
Values of the Coefficients b_n .

μ .	5.	10.	15.
b_0	0.3198	0.3685	0.3893
b_1	0.0222	0.0326	0.0247
b_2	0.0698	0.1160	0.1703

By calculating the value of the left-hand side of (11) for other values of y it was verified that the errors due to the retention of only three coefficients were small; the maximum error was equivalent to a wing-twist increasing the incidence by 2.5 per cent. over a small range for the case $\mu=15$, the errors being smaller for the other values of μ .

4. Lift and Induced Drag.

The lift L of the wing is

$$L = \rho V \int_{-R}^R K dx,$$

and the lift coefficient is given by

$$C_L = \frac{L}{\frac{1}{2} \rho V^2 \cdot 2Rc} = \frac{4\pi R\alpha}{c} \int_{-1}^1 \Gamma dy,$$

making use of the relations (4). Hence the slope of the lift curve is

$$a = \frac{dC_L}{d\alpha} = \frac{a_0}{2} \int_{-1}^1 \mu \Gamma dy, \quad (12)$$

since

$$\mu = \frac{8\pi R}{a_0 c}.$$

If we adopt Glauert's form⁽¹⁾,

$$\alpha = \alpha_0 + \frac{M}{\pi A} C_L,$$

where

$$M = \frac{\pi A}{a_0} \left(\frac{a_0}{a} - 1 \right),$$

for the angle of incidence corresponding to a given value of the lift coefficient, equation (12) gives

$$M = \frac{\mu}{4} \left[\int_{-1}^1 \frac{2}{\mu \Gamma} dy - 1 \right],$$

where $A = 2R/c$ is the aspect ratio of the wing*.

The induced drag of the wing due to the total induced velocity w is

$$D_i = \int_{-R}^R \rho w K dx.$$

Making use of (2) and (4) we obtain

$$D_i = 4\pi\rho V^2 R^2 \alpha^2 \int_{-1}^1 \Gamma(1-\mu\Gamma) dy.$$

The induced drag coefficient C_{Di} is given by

$$\frac{C_{Di}}{C_L^2} = \frac{c}{4\pi R} \frac{\int_{-1}^1 \Gamma(1-\mu\Gamma) dy}{\left[\int_{-1}^1 \Gamma dy \right]^2}.$$

Glauert's form for the induced drag coefficient is

$$C_{Di} = \frac{N}{\pi A} C_L^2,$$

and hence

$$N = \frac{\pi A C_{Di}}{C_L^2} = \frac{\int_{-1}^1 \Gamma(1-\mu\Gamma) dy}{2 \left[\int_{-1}^1 \Gamma dy \right]^2}.$$

* α_0 is the angle of incidence for two-dimensional flow which corresponds to the lift coefficient C_L , so that $a_0 = dC_L/d\alpha_0$.

The quantities M and N determined from the solutions are given in Table III.

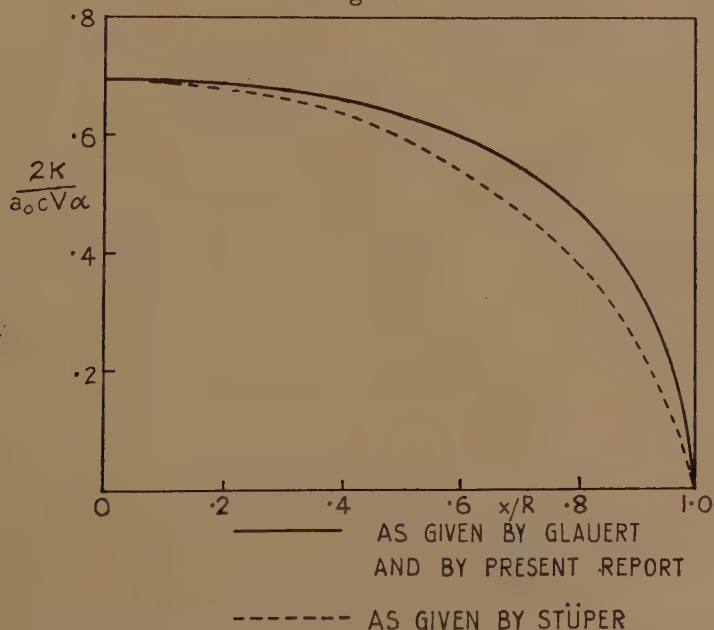
TABLE III.
Wing Characteristics.

μ .	5.	10.	15.
A/a_0	0.398	0.796	1.195
a/a_0	0.407	0.567	0.653
M	1.822	1.913	1.990
N	1.708	1.733	1.777

5. Discussion of Results.

A comparison of the distribution of circulation along the wing-span for $A/a_0=0.8$, corresponding roughly to an aspect ratio of 5, with the distributions obtained by Glauert and Stüper is shown in fig. 1; the distribution obtained is indistinguishable from Glauert's *. The values

Fig. 1.

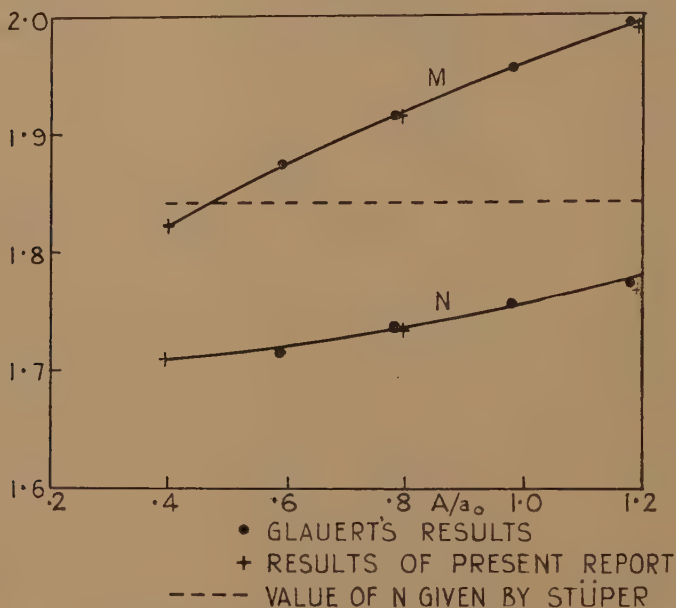


Circulation distribution for wing spanning free jet. $A/a_0=0.8$.

* This does not hold for the slopes of the distributions given by Glauert's results and by the results of the present report, which differ greatly near the wing-tips.

of M and N obtained by the different methods are plotted in fig. 2, and again agreement with Glauert's results is obtained. It may therefore

Fig. 2.



Boundary corrections for wing spanning jet.

be concluded that the error introduced in Glauert's analysis by the failure to satisfy the equation near the wing-tips is negligible.

I am indebted to Professor L. Rosenhead and to Miss H. M. Lyon for advice and criticism.

APPENDIX.

Evaluation of

$$\int_{-1}^1 \log \left(\frac{1-y'}{1+y'} \right) \left[\frac{1}{y-y'} + \frac{1}{y-1/y'} \right] dy'.$$

Let

$$\begin{aligned} I &= \int_{-1}^1 \log \left(\frac{1-y'}{1+y'} \right) \left[\frac{1}{y-y'} + \frac{1}{y-1/y'} \right] dy' \\ &= I_1 + I_2 + I_3 + I_4, \end{aligned}$$

where

$$I_1 = \int_{-1}^1 \frac{\log(1-y')}{y-y'} dy',$$

$$I_2 = - \int_{-1}^1 \frac{\log(1+y')}{y-y'} dy',$$

$$I_3 = \int_{-1}^1 \frac{\log(1-y')}{y-1/y'} dy',$$

$$I_4 = - \int_{-1}^1 \frac{\log(1+y')}{y-1/y'} dy'.$$

To determine I_1 put $y_1 = y' - y$; then

$$I_1 = - \int_{-(1+y)}^{1-y} \frac{\log(1-y-y_1)}{y_1} dy_1.$$

For $y < 0$ we write

$$\begin{aligned} I_1 &= - \int_{-(1+y)}^{(1-y)} \left[\frac{\log(1-y)}{y_1} + \frac{\log\left(1 - \frac{y_1}{1-y}\right)}{y_1} \right] dy_1 \\ &= - \int_{-(1+y)}^{(1-y)} \left[\frac{\log(1-y)}{y_1} - \sum_1^{\infty} \frac{y_1^{n-1}}{n(1-y)^n} \right] dy_1 \\ &= \left[-\log(1-y) \log|y_1| + \sum_1^{\infty} \frac{y_1^n}{n^2(1-y)^n} \right]_{-(1+y)}^{(1-y)}, \end{aligned}$$

and hence, for $y < 0$,

$$\begin{aligned} I_1 &= -\log^2(1-y) + \log(1-y) \log(1+y) \\ &\quad + \sum_1^{\infty} \frac{1}{n^2} - \sum_1^{\infty} \frac{(-)^n}{n^2} \left(\frac{1+y}{1-y} \right)^n. \end{aligned}$$

For $y > 0$ we write

$$\begin{aligned} I_1 &= - \int_{-(1-y)}^{(1-y)} \left[\frac{\log(1-y)}{y_1} + \frac{\log\left(1 - \frac{y_1}{1-y}\right)}{y_1} \right] dy_1 \\ &\quad - \int_{-(1+y)}^{-(1-y)} \left[\frac{\log|y_1|}{y_1} + \frac{\log\left(1 - \frac{1-y}{y_1}\right)}{y_1} \right] dy_1 \\ &= \sum_1^{\infty} \frac{1}{n^2} - \sum_1^{\infty} \frac{(-)^n}{n^2} - \int_{-(1+y)}^{-(1-y)} \left[\frac{\log|y_2|}{y_1} - \sum_1^{\infty} \frac{(1-y)^n}{n y_2^{n+1}} \right] dy_1 \\ &= \sum_1^{\infty} \frac{1}{n^2} - \sum_1^{\infty} \frac{(-)^n}{n^2} - \left[\frac{\log^2|y_1|}{2} + \sum_1^{\infty} \frac{(1-y)^n}{n^2 y_1^n} \right]_{-(1+y)}^{-(1-y)}; \end{aligned}$$

hence, for $y > 0$,

$$I_1 = \sum_1^{\infty} \frac{1-2(-)^n}{n^2} - \frac{1}{2} [\log^2(1-y) - \log^2(1+y)] + \sum_1^{\infty} \frac{(-)^n (1-y)^n}{n^2 (1+y)^n}.$$

In the expression for I_2 we put $y = -\bar{y}$, $y' = -\bar{y}'$, and then

$$I_2(y) = I_2(-\bar{y}) = \int_{-1}^1 \frac{\log(1-\bar{y}')}{\bar{y}-\bar{y}'} d\bar{y}' = I_1(\bar{y}) = I_1(-y).$$

Therefore, for $y > 0$,

$$I_2(y) = -\log^2(1+y) + \log(1+y) \log(1-y) + \sum_1 \frac{1}{n^2} - \sum_1 \frac{(-)^n}{n^2} \left(\frac{1-y}{1+y} \right)^n,$$

and, for $y < 0$,

$$I_2(y) = \sum_1 \frac{1-2(-)^n}{n^2} - \frac{1}{2} [\log^2(1+y) - \log^2(1-y)] + \sum_1 \frac{(-)^n}{n^2} \left(\frac{1+y}{1-y} \right)^n.$$

Adding the expressions for I_1 and I_2 we obtain, both for $y > 0$ and $y < 0$,

$$I_1 + I_2 = 4 \sum_1 \frac{1}{(2n+1)^2} - \frac{1}{2} \log^2 \left(\frac{1+y}{1-y} \right).$$

Also, since

$$\sum_1 \frac{1}{(2n+1)^2} = \frac{\pi^2}{8},$$

$$I_1 + I_2 = \frac{\pi^2}{2} - \frac{1}{2} \log^2 \left(\frac{1+y}{1-y} \right).$$

To evaluate $I_3(y)$ we put

$$y_1 = 1 - yy';$$

then

$$I_3(y) = \int_{1+y}^{1-y} \frac{1}{y^2} \left(\frac{1}{y_1} - 1 \right) \log \left(\frac{y-1+y_1}{y} \right) dy_1,$$

and hence

$$\begin{aligned} y^2 I_3(y) &= \int_{1+y}^{1-y} \frac{1}{y_1} \log \left(\frac{y-1+y_1}{y} \right) dy_1 \\ &\quad - \left[(y-1+y_1) \log(y-1+y_1) - y_1 - y_1 \log y \right]_{1+y}^{1-y} \\ &= 2y(\log 2 - 1) + \int_{1+y}^{1-y} \frac{1}{y_1} \log \left(\frac{y-1+y_1}{y} \right) dy_1, \end{aligned}$$

For $y < 0$

$$\begin{aligned} \int_{1+y}^{1-y} \frac{1}{y_1} \log \left(\frac{y-1+y_1}{y} \right) dy_1 &= \int_{1+y}^{1-y} \left[\frac{1}{y_1} \log \left(\frac{y-1}{y} \right) - \sum_1 \frac{1}{n} \frac{y_1^{n-1}}{(1-y)^n} \right] dy_1 \\ &= \log \left| \frac{y-1}{y} \right| \log \frac{1-y}{1+y} - \sum_1 \frac{1}{n^2} + \sum_1 \frac{1}{n^2} \left(\frac{1+y}{1-y} \right)^n, \end{aligned}$$

For $y > 0$

$$\begin{aligned} \int_{1+y}^{1-y} \frac{1}{y_1} \log \left(\frac{y-1+y_1}{y} \right) dy_1 &= \int_{1+y}^{1-y} \frac{1}{y_1} \left[\log \frac{y_1}{y} + \log \left(1 - \frac{1-y}{y_1} \right) \right] dy_1 \\ &= \int_{1+y}^{1-y} \left[\frac{1}{y_1} \log \frac{y_1}{y} - \sum_1 \frac{1}{n} \frac{(1-y)^n}{y_1^{n+1}} \right] dy_1 \end{aligned}$$

$$\begin{aligned}
 &= \left[\frac{1}{2} \log^2 y_1 - \log y_1 \log |y| + \sum_1^{\infty} \frac{1}{n^2} \frac{(1-y)^n}{y_1^n} \right]_{1+y}^{1-y} \\
 &= \frac{1}{2} \log^2 (1-y) - \frac{1}{2} \log^2 (1+y) - \log |y| \log \frac{1-y}{1+y} \\
 &\quad + \sum_1^{\infty} \frac{1}{n^2} - \sum_1^{\infty} \frac{1}{n^2} \left(\frac{1-y}{1+y} \right)^n.
 \end{aligned}$$

Hence, for $y < 0$,

$$y^2 I_3(y) = 2y (\log 2 - 1) + \log \left| \frac{y-1}{y} \right| \log \frac{1-y}{1+y} - \sum_1^{\infty} \frac{1}{n^2} + \sum_1^{\infty} \frac{1}{n^2} \left(\frac{1+y}{1-y} \right)^n,$$

and, for $y > 0$,

$$\begin{aligned}
 y^2 I_3(y) &= 2y (\log 2 - 1) + \frac{1}{2} \log^2 (1-y) - \frac{1}{2} \log^2 (1+y) \\
 &\quad - \log |y| \log \frac{1-y}{1+y} + \sum_1^{\infty} \frac{1}{n^2} - \sum_1^{\infty} \frac{1}{n^2} \left(\frac{1-y}{1+y} \right)^n.
 \end{aligned}$$

To determine I_4 we substitute $y = -\bar{y}$, $y' = -\bar{y}'$, and obtain

$$I_4(y) = I_4(-\bar{y}) = \int_{-1}^1 \frac{\log (1-\bar{y}')}{\bar{y} - 1/\bar{y}'} d\bar{y}' = I_3(\bar{y}) = I_3(-y)$$

Hence, for $y < 0$,

$$\begin{aligned}
 y^2 I_4(y) &= -2y (\log 2 - 1) + \frac{1}{2} \log^2 (1+y) - \frac{1}{2} \log^2 (1-y) \\
 &\quad - \log |y| \log \frac{1+y}{1-y} + \sum_1^{\infty} \frac{1}{n^2} - \sum_1^{\infty} \frac{1}{n^2} \left(\frac{1+y}{1-y} \right)^n,
 \end{aligned}$$

and, for $y > 0$,

$$y^2 I_4(y) = -2y (\log 2 - 1) + \log \left| \frac{1+y}{y} \right| \log \frac{1+y}{1-y} - \sum_1^{\infty} \frac{1}{n^2} + \sum_1^{\infty} \frac{1}{n^2} \left(\frac{1-y}{1+y} \right)^n.$$

Hence, for $y > 0$ and $y < 0$,

$$y^2 (I_3 + I_4) = \frac{1}{2} \log^2 \frac{1-y}{1+y}.$$

Adding these results together we obtain

$$I = \frac{\pi^2}{2} + \frac{1-y^2}{2y^2} \log^2 \left(\frac{1-y}{1+y} \right).$$

References.

- (1) Glauert, Aero. Research Committee R. & M. 1603 (1934).
- (2) Stüper, 'Ingenieur-Archiv,' 1932, p. 338.
- (3) Stüper, *Luftfahrtforschung*, vol. xii. p. 267 (1935).
- (4) Diprose, Aero. Research Committee R. & M. 1813 (1937).
- (5) Lotz and Riegels, *Luftfahrtforschung*, vol. xiii. p. 550 (1937).
- (6) Betz and Petersohn, *Zeitsch. angew. Math. Mech.* viii. p. 253 (1928).

XXIII. *On the Development of Turbulent Liquid Motion
over an Infinite Plate.*

By MANOHAR RAY, Lahore, India *.

[Received August 19, 1938.]

1. *Introduction.*

THERE exist at present two theories of turbulence based on the idea of the transport of momentum by Prandtl ⁽¹⁾ and that of vorticity by Taylor ⁽²⁾. The equations have, however, been solved for only a few cases. In every case a comparison of the solutions of the problem by the two theories with observation has furnished a basis for testing the truth or otherwise of each theory. In the present paper a problem on the growth of turbulence on a very large plate, set up by some external agency for example, has been studied. The flow is supposed to be two dimensional, and though depending on time, is taken to be uniform in every section across the layer of turbulence. The conclusions will relate to the spread of a two dimensional turbulent motion perpendicular to a plate, while the motion at any instant in every section perpendicular to the direction of flow remains identical. Outside of the turbulent layer the motion is uniform. There are two limitations of the results obtained. Full account has not been taken of the gradual transition of a boundary layer where viscosity is predominant to another where turbulence plays the prominent role through an intermediate layer where both effects are comparable. Hence the solutions cannot be expected to give much correct information very near the plate. But the viscous effect has been to some extent taken into consideration by calculating the total depth of the layers according to the one-seventh law of Blasius. Secondly, on account of the *ad hoc* assumption of the existence of turbulent condition, which should be the same in all sections perpendicular to the direction of flow at any instant of time, our results cannot be extrapolated backward to the initial stage of turbulence, so that the results for $t=0$ are not also expected to have a meaning.

It has been found that Taylor's equation, based on the idea of transport of vorticity, can be solved yielding certain definite results, but a similar successful attempt is not possible for Prandtl's equation for the transport of momentum.

* Communicated by the Author.

2. Field Equations and their Solutions.

Take the origin on the plate, the axis of x parallel and that of y perpendicular to the plate. The motion in all sections perpendicular to the direction of flow is supposed to be similar. The field velocity of flow at any instant is a function of y only, and there is no velocity in the y -direction. By field velocity is meant the velocity obtained by subtracting the turbulent fluctuations from the actual velocity. Thus we take

$$u=u(y, t), v=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The equations of turbulent motion on the momentum ⁽¹⁾ and vorticity ⁽²⁾ transport theories are

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial y} \left[l^2 \left(\frac{\partial u}{\partial y} \right)^2 \right] \quad . \quad . \quad . \quad . \quad . \quad . \quad (2 \text{ P})$$

and
$$\frac{\partial u}{\partial t} = l^2 \frac{\partial u}{\partial y} \frac{\partial^2 u}{\partial y^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2 \text{ T})$$

respectively, l being the mixture length. The effect of liquid pressure is here neglected in comparison with turbulence stresses.

To effect a solution we make the following special assumptions for the velocity and the mixture length,

$$u=u_0 f(\eta), \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where the non-dimensional quantity η is defined by

$$\eta = \frac{y}{\phi(t)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and
$$l^2 = \alpha y^2 \chi(t). \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Here u_0 is the velocity of flow in the main stream, $\phi(t)$ and $\chi(t)$ are functions of t only, and α is a constant. The assumption (5) is consistent with the idea that the mixture length tends to the limit zero on the plate.

Making the substitutions (3) and (5) in the equations (2 P) and (2 T) we have

$$2 \alpha u_0 \frac{\chi(t)}{\phi(t)} f' (f' + \eta f'') = -f' \frac{\phi'(t)}{\phi(t)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (6 \text{ P})$$

and
$$\alpha u_0 \frac{\chi(t)}{\phi(t)} \eta f' f'' = -f' \frac{\phi'(t)}{\phi(t)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (6 \text{ T})$$

where dashes denote differentiations with respect to corresponding arguments.

Both equations give firstly

$$\frac{\chi(t)}{\phi(t)} = \frac{\phi'(t)}{\phi(t)},$$

that is,

$$\chi(t) = \phi'(t). \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Neglecting the solution $f' = 0$, which gives $u = \text{constant}$, we get further from (6 P) and (6 T)

$$f' + \eta f'' = -\frac{1}{2\alpha u_0}, \quad \dots \dots \dots (8 \text{ P})$$

and

$$f'' = -\frac{1}{\alpha u_0 \eta} \dots \dots \dots (8 \text{ T})$$

In each case the equation can be completely integrated. Equation (8 P) has the integral

$$f = c\eta + d \log \eta + e, \quad \dots \dots \dots (9 \text{ P})$$

where $c = -\frac{1}{2\alpha u_0}$, and d and e are constants with respect to η . Similarly equation (8 T) gives on integration

$$f = c_1 \eta \log \eta + d_1 \eta + e_1, \quad \dots \dots \dots (9 \text{ T})$$

where $c_1 = -\frac{1}{\alpha u_0}$, and d_1, e_1 are constants of integration.

The solutions of (2 P) and (2 T) have thus the forms

$$u = u_0(c\eta + d \log \eta + e), \quad \dots \dots \dots (10 \text{ P})$$

and

$$u = u_0(c_1 \eta \log \eta + d_1 \eta + e_1), \quad \dots \dots \dots (10 \text{ T})$$

or replacing η by (4),

$$u = u_0 \left\{ c \frac{y}{\phi(t)} + d \log \frac{y}{\phi(t)} + e \right\}, \quad \dots \dots \dots (11 \text{ P})$$

and

$$u = u_0 \left\{ c_1 \frac{y}{\phi(t)} \log \frac{y}{\phi(t)} + d_1 \frac{y}{\phi(t)} + e_1 \right\}. \quad \dots \dots \dots (11 \text{ T})$$

We thus express the velocity u in terms of *three* constants, and besides there is the function $\phi(t)$ to be determined.

These solutions cannot be expected to be valid in the immediate neighbourhood of the wall where viscosity predominates. In this layer the shear is great and the velocity rises very quickly from zero at the surface of the plate, so that $\frac{\partial u}{\partial y}$ is very large. Next to this layer there exists another in which both viscosity and turbulence are of importance. Our equations (2 P), (2 T), however, relate to a layer where turbulence predominates. In order to get the proper orientation for our problem we shall neglect the finite widths of the first two layers and assume that $\frac{\partial u}{\partial y} = \infty$ is the condition to be satisfied on the plate. On the plate surface $y = 0$, neither of the solutions (11 P) and (11 T) gives a zero value for u . The vorticity transport theory gives a finite value, while the momentum theory gives a negative infinite value. But both theories

give a positive infinite value for $\frac{\partial u}{\partial y}$ at the wall, the apparent negative sign in the case of the vorticity theory is changed to positive by the sign of c_1 , which is itself negative.

We have another condition to be satisfied, namely the turbulent flow should ultimately merge into the constant flow u_0 . But a condition like $u \rightarrow u_0$ as $y \rightarrow \infty$ cannot be satisfied by either of the solutions (11 P) and (11 T). We look, therefore, for a sharply defined boundary layer of finite thickness defined by $u = u_0$, $\frac{\partial u}{\partial y} = 0$ for $y = y_0$. These conditions, however, do not fully determine the three constants either in (11 P) or in (11 T). Further, we are required to find the function $\phi(t)$. This can be calculated from a knowledge of the variation of y_0 , the width of the boundary layer with the time t , whilst the value y_0 of y for which $u = u_0$ and $\frac{\partial u}{\partial y} = 0$, and another value y_1 of y for which u/u_0 is given, are sufficient to determine the three constants. We first find out the variation of y_0 with t . This is not directly given by our equations, but can be obtained by Karman's Integral-condition ⁽³⁾ for the boundary layer, namely,

$$\frac{\partial}{\partial t} \int_0^{y_0} \rho u dy + \frac{\partial}{\partial x} \int_0^{y_0} \rho u^2 dy - u_0 \frac{\partial}{\partial x} \int_0^{y_0} \rho u dy = -y_0 \frac{\partial p}{\partial x} - \tau_0, \quad (12)$$

which in the present case reduces to

$$\frac{\partial}{\partial t} \int_0^{y_0} \rho u dy = -\tau_0. \quad (13)$$

For u and τ_0 we take here the forms given by Blasius for a turbulent layer, namely

$$u = u_0 \left(\frac{y}{y_0} \right)^{1/7} \quad (14)$$

and

$$\tau_0 = 0.023 \rho u_0^2 \left(\frac{\nu}{u_0 y_0} \right)^{1/4}, \quad (15)$$

ν being kinematic viscosity.

Making use of (14) and (15) in (13), we get at once

$$\frac{dy_0}{dt} = -0.023 \times \frac{8}{7} u_0 \left(\frac{\nu}{u_0 y_0} \right)^{1/4}, \quad (16)$$

which on integration gives

$$y_0 = (0.0329)^{4/5} \left(\frac{\nu}{u_0^2 t} \right)^{1/5} u_0 t. \quad (17)$$

Comparing this with (4) we put

$$\phi(t) = \left(\frac{\nu}{u_0^2 t} \right)^{1/5} u_0 t, \quad (18)$$

suppressing a numerical constant factor, which only serves as a magnification factor (the effect of which is only to magnify uniformly the value of y_0 in Table II.).

Consequently
$$\chi(t) \sim 1/t^{1/5}, \quad \dots \dots \dots (19)$$

and
$$y_0 = (.0329)^{4/5} \phi(t) \quad \dots \dots \dots (20)$$

3. Discussion of the Solution.

We now discuss in detail the solution (11 T) as given by the Taylor's vorticity transport theory. c_1, d_1, e_1 in (11 T) are absolute constants. This solution has to fit the conditions at the boundaries. We have already seen this solution gives infinite value for $\frac{\partial u}{\partial y}$ on $y=0$. Further we have the upper boundary conditions, namely for some $y=y_0(t)$, $u=u_0$ and $\frac{\partial u}{\partial y}=0$. These enable us to determine d_1 and e_1 in terms of c_1 which is an absolute constant, being equal to $-\frac{1}{\alpha u_0}$. The two conditions at the upper boundary give

$$c_1 \frac{y_0}{\phi(t)} \log \frac{y_0}{\phi(t)} + d_1 \frac{y_0}{\phi(t)} + e_1 = 1,$$

and
$$c_1 \log \frac{y_0}{\phi(t)} + c_1 + d_1 = 0;$$

whence
$$d_1 = -c_1 \{ 1 + \log (y_0/\phi(t)) \}$$

and
$$e_1 = 1 + c_1 y_0/\phi(t).$$
 (21)

We note that $y_0/\phi(t)$ is a numerical constant as given by (20). Substituting the values of d_1 and e_1 from (21) in (11 T) we get the velocity distribution for the vorticity transport theory as follows :

$$\frac{u_0 - u}{u_0} = -c_1 \frac{y_0}{\phi(t)} \left\{ 1 - \left(\frac{y}{y_0} \right) + \left(\frac{y}{y_0} \right) \log \left(\frac{y}{y_0} \right) \right\} \quad \dots \quad (22)$$

We note here that c_1 is a dimensionless constant, for from (18) and (7) we see that $\chi(t)$ is of dimension u_0 , so that from (5) α must be of dimension $1/u_0$, i. e., αu_0 is a non-dimensional quantity, i. e., c_1 is a numerical constant. The value of this constant c_1 can be determined if we know the value of u/u_0 at some point within the boundary layer at any given instant.

We have seen that the factor $-c_1 y_0/\phi(t)$ is as a whole a numerical one, so that apart from the magnification factor $-c_1 y_0/\phi(t)$ the profile of $(u_0 - u)/u_0$ can be represented by the curve obtained by plotting $\{1 - (y/y_0) + (y/y_0) \log (y/y_0)\}$. Two cases have been considered. First of all, the Table I. and fig. 1 represent the profile $(u_0 - u)/u_0$ at any time for different

values of (y/y_0) , *i. e.*, at points whose depths are definite fractions of the (varying) total thickness of the turbulent layer. Secondly, to trace the change of velocity with time at *any given point within the turbulent layer*, we put

$$(.0329)^{4/5} \left(\frac{\nu}{u_0^2} \right)^{1/5} u_0 = 1. \quad . \quad . \quad . \quad . \quad . \quad (23)$$

This equation, together with the numerical values of t as given in Table II., only fixes our units of length and time, which are here measured on a new scale. With the help of (23), (17) gives

$$y_0 = t^{4/5}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

Results.

TABLE I.

y/y_0	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
$(u_0 - u)/u_0$	1	.6697	.4782	.3388	.2335	.1534	.0935	.0503	.0217	.0051	0

TABLE II.

t	y_0	$\left(\frac{u_0 - u}{u_0} \right) \left(\frac{y}{y_0} \right)_{t=1} = .1$	$\left(\frac{u_0 - u}{u_0} \right) \left(\frac{y}{y_0} \right)_{t=1} = .5$	$\left(\frac{u_0 - u}{u_0} \right) \left(\frac{y}{y_0} \right)_{t=1} = .9$
1	1	.6697	.1534	.00516
2	1.741	.7785	.3545	.142
3	2.408	.8263	.466	.2584
4	3.0315	.8545	.5378	.3426
5	3.6239	.8733	.5888	.4057
6	4.193	.8871	.6271	.4551
7	4.7436	.8976	.6574	.495
8	5.278	.9059	.6821	.5279
9	5.7995	.9127	.7025	.5557
10	6.3096	.9185	.72	.5797

Hence, if we start with $t=1$, initially $y_0=1$, *i. e.*, the thickness of the turbulent layer is unity. We take the point y where we seek to trace the change of velocity with time to lie within this unit thickness. Three cases have here been considered, namely

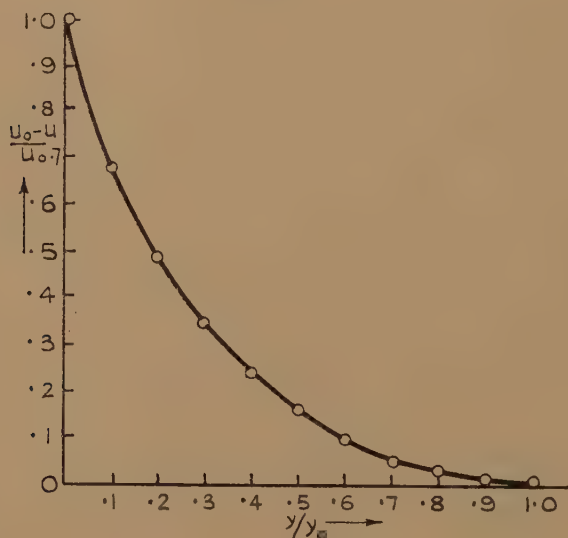
$$\left(\frac{y}{y_0} \right)_{t=1} = .5, \quad \left(\frac{y}{y_0} \right)_{t=1} = .1, \quad \text{and} \quad \left(\frac{y}{y_0} \right)_{t=1} = .9.$$

In the first case, for example, y is given the fixed value .5. Then y_0 is calculated from (24) for different values of the time t . From (22) $(u_0 - u)/u_0$ may now be computed apart from a constant magnifying factor (since $c_1 y_0 / \phi(t)$ is a numerical constant) for different values of the time. This would give the values of the velocity at a point midway in the initial width of the turbulent layer at different times. Table II.

and fig. 2 show these changes in velocities at three points within the layer.

It is easily seen from fig. 1 that if we divide the turbulent layer at any instant of time into four equal parts, then within the first quarter from the top the velocity falls from u_0 only about 3 per cent., while to the middle of the layer and to the ends of the third quarter the falls in velocities are 15 and 40 per cent. respectively. It is only within the last quarter near the plate that the fall is most rapid, amounting to about 60 per cent.

Fig. 1.



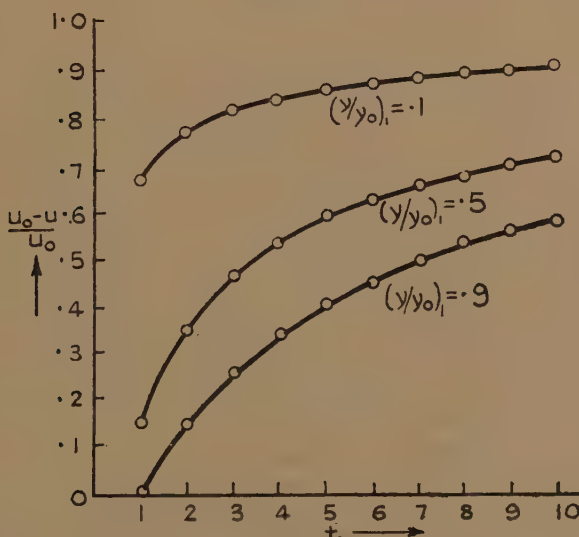
Shows the drop in velocity in the turbulent layer at different depths. y_0 is the thickness of the layer at any instant of time, and y the distance of a point in the layer from the plate. The curve shows the drop in velocity in moving through the layer from the top in terms of y/y_0 ; this velocity profile is independent of the time.

of the total fall. An examination of fig. 2 shows that at a fixed point (which may be observed) very near to the top of the turbulent layer at the initial time $t=1$, the rise in velocity with time is almost linear for a certain time, and as this point of observation is taken more and more within the layer at $t=1$, this linear rise gradually disappears. . At a point near the plate the velocity approaches a constant value at a quick rate at first and then very slowly. The course of velocity at three points (1) near the top of the layer, (2) near the plate, and (3) at the middle, at $t=1$, is shown in fig. 2.

We may notice that while the vorticity transport theory of Taylor gives definite results for the present problem, the alternative momentum transport theory leads to solutions giving infinite value for the velocity at the plate, and so cannot be expected to give correct information as to the distribution of velocity within the layer.

It is necessary to restate the limitations of the above solution explained in the introduction. It is not expected to give correct results very near the plate, nor can it be extrapolated to $t=0$, the initial stage of turbulence.

Fig. 2.



Shows the change in velocity $(u_0 - u)/u_0$, with time at a fixed point within the layer with the growth of turbulence. The scale of time t has been chosen suitably. The values $(y/y_0) = 0.1, 0.5, 0.9$, determine the positions of three points within the turbulent layer at time $t=1$. Note that the drop in velocity given by $(u_0 - u)/u_0$ is more rapid at points further from the plate than at points which lie nearer.

4. Problem of Temperature Distribution.

The associated problem of heat conduction can also be treated, though not quite so satisfactorily. If the plate be heated and if θ be the difference of the temperatures between the plate and the fluid, the differential equation for θ is

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial y} \left(l^2 \frac{\partial u}{\partial y} \frac{\partial \theta}{\partial y} \right), \quad \dots \dots \dots (25)$$

where l is the same as in (5).

If we put $\theta = \theta_0 g(\eta)$, we find from (32) with the help of (3), (4), (5), and (8 T), that

$$2g' + \eta g'' = 0, \quad \dots \dots \dots (26)$$

dashes meaning differentiation with respect to η .

On integration we get

$$\frac{\theta}{\theta_0} = g = A - \frac{B}{\eta}, \quad \dots \dots \dots (27)$$

where A and B are constants and θ_0 is the temperature difference between the fluid at infinity and the plate.

The equation (27) gives θ in terms of y and t :

$$\frac{\theta}{\theta_0} = A - \frac{B\phi(t)}{y} \dots \dots \dots (28)$$

If we assume that on the boundary y_0 , $\theta = \theta_0$, we get

$$A = 1 + \frac{B\phi(t)}{y_0}, \quad \dots \dots \dots (29)$$

hence
$$\frac{\theta_0 - \theta}{\theta_0} = \frac{B\phi(t)}{y_0} \left(\frac{y_0}{y} - 1 \right) \dots \dots \dots (30)$$

Moreover, A and B being absolute constants, $B\phi(t)/y_0$ is a purely numerical factor. The profile of $(\theta_0 - \theta)/\theta_0$, apart from a magnification factor, is given by $(y_0/y - 1)$, and is consequently a straight line.

It is to be noted that (30) cannot give the correct distribution of temperature near the plate, but may correctly represent the conditions near the upper boundary.

In conclusion I want to express my thanks to Prof. N. R. Sen, of the Calcutta University, for kind help in this work.

References.

- (1) Prandtl, *Zamm*, v. p. 136 (1925).
- (2) Taylor, *Proc. Roy. Soc. A*, cxxxv. p. 685 (1932).
- (3) v. Karman, *Zamm*, i. p. 235 (1921).

Forman Christian College,
Lahore, India.

XXIV. *Geiger Point Counters.*

By B. DASANNACHARYA, M.A., Dr.Phil., F.Inst.P., Professor of
Physics, Benares Hindu University, Benares, India,
and AMAR CHAND SETH, M.Sc. (Benares) *.

[Received August 29, 1938.]

1. *Introduction.*

It was shown by B. Dasannacharya and T. S. Krishna Moorthy † that the number of counts recorded with a point counter in the region between the threshold voltage and the voltage commencing the region of saturation counts, shows a marked diminution along the axis of the counter as compared with the counts on either side at right angles to the axis. This diminution or dip is very marked when the voltage is low and near the threshold value and becomes inappreciable and entirely disappears as the saturation range of voltage is reached.

The present paper gives an account of further experiments carried out to explain the phenomenon. It is found that *the dip is due to the inability of the counter to count more than a certain number of particles at a given voltage*. If the particles exceed this number, the counter ceases to amplify and the counts are missed. If the leakage of charge from the central needle is facilitated by lowering the leak resistance the dip may become less prominent or even entirely disappear.

A phenomenon similar to the above can be obtained if the resistance connecting the plate and the battery is very high and can be explained in a similar way. To the same cause must be attributed the fall in the number of counts when a radium source is moved towards a counter along the axis. The counts begin to increase, but will diminish when the source is so near the counter and the number of ionizing particles so large that the counter cannot magnify.

Fig. 1 (a) shows a sketch of the connexions, and fig. 1 (b) the characteristic of a counter. Let the number of ionizing particles that enter the counter be large, then the voltage of the needle rises by an amount

$$v_c = iR_c$$

* Communicated by the Authors.

† B. Dasannacharya and T. S. Krishna Moorthy, Phil. Mag. (7) xxiii. p. 609 (1937).

and the voltage of the plate falls by an amount equal to

$$v_b = iR_b,$$

where R_b and R_c are the resistances connecting the plate to the battery and the resistance connecting the needle to the earth, respectively; i is the current passing through the counter.

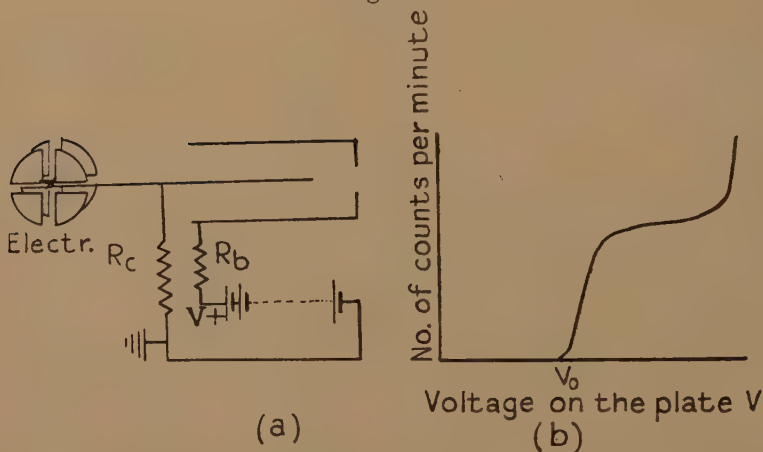
The condition that the particles will cease to be counted will be obtained when

$$V - (v_b + v_c) < V_0;$$

here V_0 is the voltage at which counting will just start, namely, the threshold voltage, and V is the voltage of the battery (see fig. 1). If

$$\Delta V = V - (v_b + v_c) - V_0,$$

Fig. 1.



then the counts will be registered with an amplitude depending on ΔV . Now both v_b and v_c will fluctuate about a mean value on account of the number of particles or ionizing sources entering the counter not being so many that a continuous current will pass. If the number so entering is so small that a particle does not come before the voltage set up by the previous count has died out, *i. e.*, when $v_b = v_c = 0$, then the full amplitude will be obtained for all the counts and the counts would not be missed. From this it would follow that the counting capacity will rise with increase of voltage, but diminish with an increase in the number of counts. So if saturation counting region is to be obtained it is subject to the condition that the number to be counted and which enters into the counter is not too many. For theoretical considerations of discharge see the paper by v. Hippel* and Sven Werner†, the latter, however, dealing with line

* v. Hippel, *Zeit. d. Physik*, xcvii. p. 455 (1935).

† S. Werner, *Zeit. d. Physik*, xc. p. 384 (1934); xcii. p. 705 (1934).

counters only. The following is therefore a contribution to the role of resistance in the Geiger point counters.

2. *Experimental Arrangement.*

The arrangement remained practically the same as was used by Dasannacharya and Krishna Moorthy (*loc. cit.*). In some of the experiments the old radon tube, which was used as the source of radiation and which consisted of about four thin-walled capillary tubes, has been used, and in some only one tube without any lead shield was employed, particularly in the experiments in sections 5 and 6. The resistance in the circuit connecting the battery to plate of the counter used to be a water resistance of about half a megohm; in the present series arrangement has been made to vary this resistance within very wide limits, even ten thousand megohms could be put in. The xylol alcohol resistance was of the sliding type, as was the resistance formerly used to connect the central needle of the counter to the earth. According to the range required the xylol-alcohol mixture was varied.

3. *The Axial Dip at different Voltages and Pressures.*

The axial dip was obtained by Dasannacharya and Krishna Moorthy at different voltages, but the pressure of the air in the counter was constant, namely, 3.0 cm. Hg.

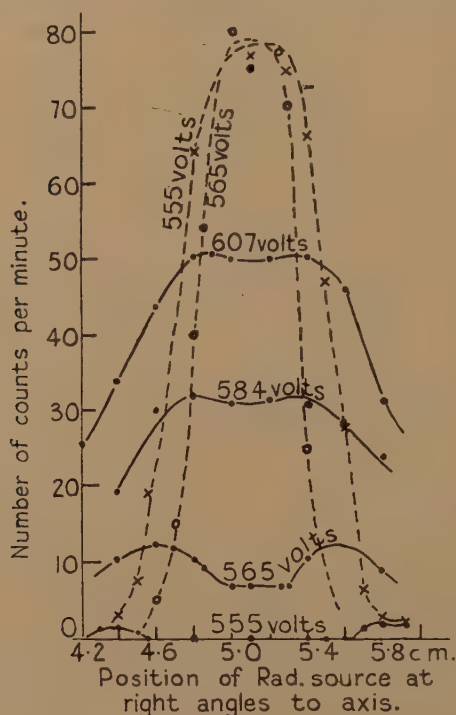
Fig. 2 shows curves for the axial dips for a pressure of 1.58 cm. for four different voltages. The counts at any position on the line perpendicular to the axis and marked as abscissæ, are shown as ordinates. A feature of those curves, not found in fig. 3 of Dasannacharya and Krishna Moorthy, is that at the low voltages, namely, of 555 volts, the axial dip becomes actually a zero count region extending from 4.56 cm. up to 5.65 cm., *i. e.*, extending to nearly 5.5 mm. on either side of the axis. With 565 volts the axial dip is not a zero region. The dip becomes insignificant at the higher voltages of 584 and 607.

In fig. 2 there are two other curves shown in dashed lines corresponding to voltages of 555 and 565. They give the number of counts which are not amplified to any great extent. These small amplitude counts, amplitude less than about one-twentieth of the amplitude of the magnified counts shown in the other curves of the figure, first become noticeable about the positions on the perpendiculars to the axis which mark the maximum for the big counts. The small amplitude counts are particularly large when the big counts diminish and the position of the axial dip attained. On the axis of the counter the small amplitude counts are largest in number. At higher voltages, where the dip is getting more and more insignificant, the number of small amplitude counts also tend to become increasingly insignificant. There is no doubt, consequently,

that the dip is connected with the existence of the small amplitude counts. As explained in section 1, it is due to the too large a number of ionizing agencies entering the counter, more than can be amplified.

Fig. 3 is for a pressure of 1.0 cm. of Hg. The amplified counts curve for 515, 537, and 555 volts are shown. The small amplitude curves for voltages of 515 and 555 are also shown. The general features of the highly amplified curves are the same as those of fig. 2. The small amplitude curves show clearly that the higher the voltage the less is the

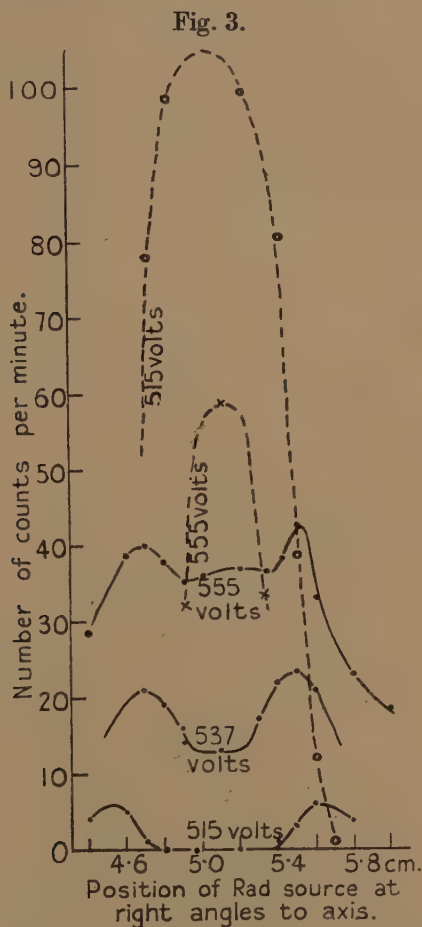
Fig. 2.



number of small amplification counts. It may be remarked that for very low voltages even the small amplitude curves are unobservable.

Curves similar to the above have been observed for pressures of 0.4 and 2.4 cm. While the features referred to above are the same, the disparity between the large and small amplitudes becomes less and less as pressure is lowered, in that large amplitudes become less and small become greater. The big amplitude at 0.4 cm. pressure was 26 mm. and small 4 mm. of the scale in the eye-piece of the microscope reading the deflexions in the Lindemann electrometer connected to the central needle.

The corresponding amplitudes at 2.4 cm. pressure were 65 and one, at one of the lower voltages.

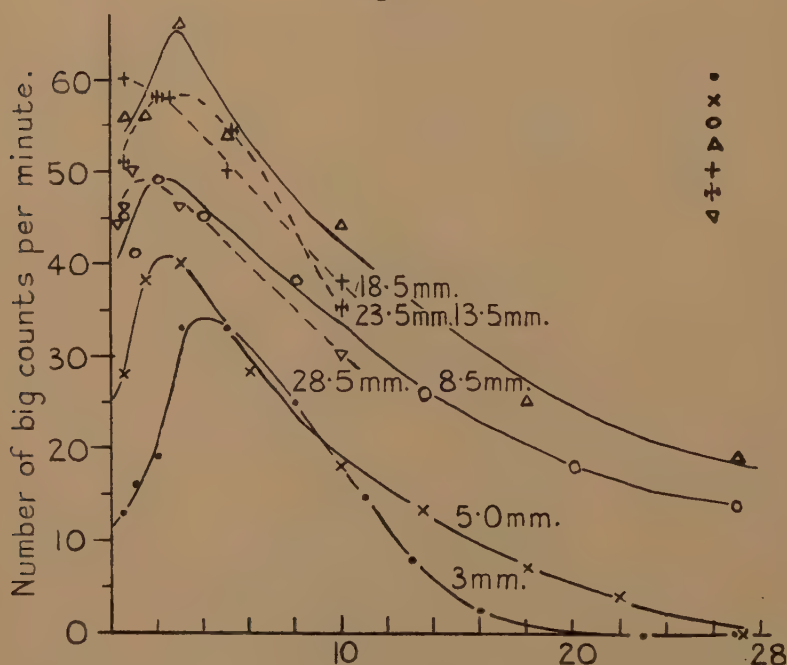


4. *The Dependence of Counts on the Leak Resistance R_c connecting the Needle Electrode to the Earth.*

In fig. 4 the abscissæ give the values of R_c and the ordinates give the number of big amplitude counts. The number on any one curve denotes the distance of the radiating source on the axis measured from the front surface of the window of the counter. This position remains the same for the readings of any one curve. The voltage on the plate was 516. The resistance R_b connecting the plate to the battery was common to all the curves of this figure, and was 36 megohms. The air pressure in the counter was 1.5 cm.

In the curve marked for a distance of source equal to 3 mm., the count for $R_c = 28$ ($1 = 200 \cdot 10^6$ ohms, is zero. At 18 instead of 28, on the alcohol xylol mixture it is 1 per minute, at 13 it is 8, reaches a maximum of 34.5 counts per minute at $R_c = 4$; the counts corresponding to still smaller values for R_c fall down. It is 13 at 0.5 for R_c . The other curves are for greater distances of the source on the axis of the counter. For greater distance, the number of ionizing agencies entering are fewer; hence the number of counts are larger for reasons explained in section 1. For very small values of R_c the counts are fewer, *i. e.*, there is a dip in

Fig. 4.



Leak resistance R_c , 1 div = 200 megohms from needle to earth.

Numbers on the curves give the distance of the radium needle source from the window along the axis in mm. Resistance from battery to counter, $R_b = 36$ megohms.

the curve. It is because the current passing through, due to the large number of ionizing agencies, is so large that the voltage drop across the R_b reduces the effective plate voltage, with the consequence that counts are not magnified, hence missed.

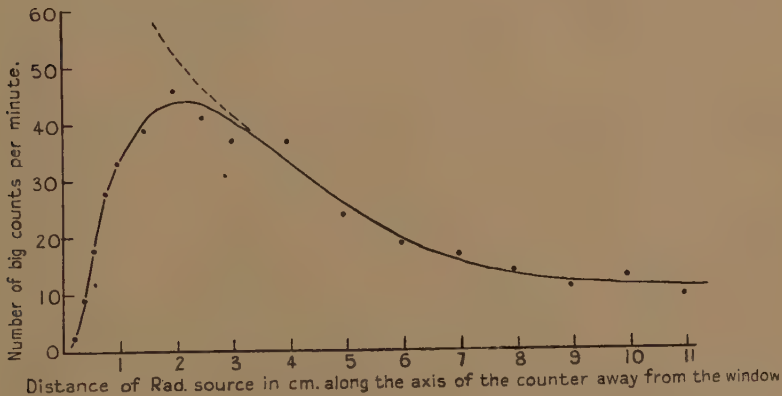
The largest record in the counts takes place for a position of source on axis equal to 13.5 mm. from the window. It is 66 per minute for R_c

equal to 3. For further increase in distance of the source the counts for any one value of R_c go on diminishing. We thus see how the number of counts recorded depend on the value of the leak resistance.

5. *Dependence of Counts on the Position of the Source on the Axis of the Counter.*

If in fig. 4 one were to draw a line parallel to the axis of counts, from any point on the R_c axis, we get the counts recordable for constant and given value of R_b , and constant values of R_c , which, however, can be varied between wide limits. One such curve is shown in fig. 5 taken for another setting of R_c and R_b than that shown in fig. 4. Very near to the

Fig. 5.



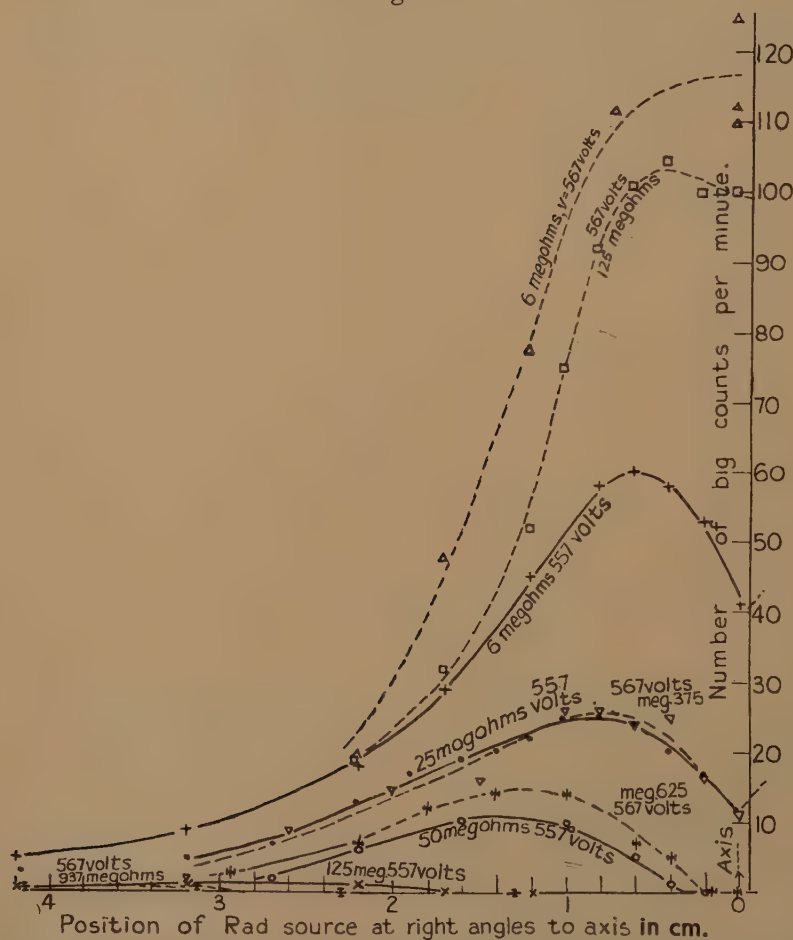
counter for the source, fig. 5 shows that the counts are very few. The number rises to a maximum of 45 counts per minute at a distance of 20 mm., and for further increase in distance continues to fall off.

The dotted line in fig. 5 shows the number of counts to be expected if the inverse square law were to hold good on the basis of the number of counts obtained at 10 cm. We see that the inverse square law holds good until the source is brought down to 3.5 cm. on the axis. For smaller distances the ionizing agents are too many for the full number to be recorded. The number of counts missed or not properly amplified will be given roughly by the deviation of the curve from the dotted inverse square curve. With source at 2 cm. another source of radium of eleven milligrams strength was brought near and the big counts entirely disappeared. This shows very clearly how the counting capacity diminishes with the number to be counted if such number exceeds a given value dependent on the voltage of the plate.

6. Dependence of Counts on the Plate Resistance R .

The arrangement to show this is the same as that in section 4 used to study dip with variations of R_c , the leak from the central needle electrode, keeping R_b small 36 megohms, and constant. Here in the present case, the value of R_c is kept at about 150 megohms, the same in all the trials

Fig. 6.



in this section. The dip is studied for 557 and 567 volts, the pressure in the counter remaining at 1.5 cm. (see fig. 6).

At 557 volts with R_b equal to 10 mm., *i. e.*, 125 megohms, the points for counts are marked with \times , dip extends to about 1.5 cm. on either side of the axis. For sparing space counts only to one side of the axis are shown, as the curve on the other side will be quite symmetrical with

this. With R_b as 4 mm., *i. e.*, 50 megohms, the points are marked with circles, the counts reach a maximum at about 1.4 cm. from the axis, and the zero dip occurs at only 3.5 mm. from the axis. With 2 mm. the maximum number of counts are 25 per minute, occurs at 1 cm. from the axis, the dip on the axis is no more zero, but is about 10 per minute. At half mm. resistance, 6 megohms, the maximum of counts 60 per minute is attained at a distance of 9 mm. from the axis, and the dip value is 37 on the axis.

At the higher voltage of 567 the same general features are evident. The dip has completely disappeared at 0.5 mm. recording 120 counts per minute. The dip is noticeable at 10 mm., the maximum counts being attained at 5 mm. away from axis, with 104 counts per minute, the dip value being 90 per minute. With 30 mm. for R_b the maximum value, 26 per minute is attained at 9 mm. from the axis, and the dip value on the axis is 10 per minute. With 50 mm. of R_b the zero dip region extends to a considerable distance from the axis. The same observation holds for 75 mm. for R_b .

These show very clearly how in investigations using counters to detect ionizing agencies the resistances in the different parts of the counter play an important role.

7. Summary and Conclusion.

Section 1 shows the different possibilities whereby the functioning of a point counter can be influenced by the resistances in the counter circuits. In 2 the alterations in the arrangements for observations of counts not mentioned in the investigation by Dasannacharya and Krishna Moorthy are given. In 3 the axial dip caused by the failure of a counter due to the large leak resistances connecting the central needle electrode to the earth, to amplify the current, is investigated for pressures of 1.58 cm., 1.0 cm., 0.4 cm., and 2.4 cm., but the curves are given in figs. 2 and 3 only for the first two pressures mentioned. The curves for the less amplified counts are also given, and it is shown that the dip is connected with the number of unamplified counts caused by insufficiency of amplification.

In sections 4 and 5 it is shown that the leak resistance R_c brings about vast changes in the number of counts recorded, depending also on the number of ionizing agencies entering the counter got either by moving the radiation source away from the axis at right angles or away from the counter along the axis. Finally, the role of the resistance R_b connecting the battery to the plate is shown to be the same as that of R_c .

Department of Physics,
Benares Hindu University,
Benares, India.
August 22nd, 1938.

XXV. *Notices respecting New Books.*

A Textbook of Thermodynamics. By F. E. HOARE. Second edition. [Pp. xii+307.] (London: Edward Arnold & Co., 1938. Price 15s.)

THE comparatively early appearance of a second edition of this book is a welcome sign of the wide appreciation which it has received. The author is concerned to show thermodynamics in action, and after the first few chapters dealing with basic principles and the more direct applications, he discusses, with particular reference to the bearing of thermodynamics, a wide variety of topics. There are chapters on equations of state, change of phase, gaseous equilibrium, liquids and solutions, electrical phenomena, radiation, and specific heat. The mathematical treatment is straightforward, aiming at fullness of statement and development from first principles, rather than at elegant conciseness which often defeats its object.

The second edition has been much improved by the consistent adoption of the symbolism recently recommended by a Joint Committee of the Chemical, Faraday, and Physical Societies. A good collection of examples has been added, an unusually high proportion of which have point and practical interest to add to the challenge to skill in formal manipulation.

Critical comment on a few points, for consideration in connexion with future editions, may not be out of place in a book whose general character is so admirably adapted to its purpose, and which is likely to be widely used. The avoidance of the use of the partial differentiation symbol seems misguided in a field where the clear distinction of ordinary and partial differentiation is so advantageous. The formulation of Nernst's heat theorem in a manner implying an absolute measure of entropy rather than of entropy differences is hardly defensible in a purely thermodynamic context, and it could be avoided without detriment to any of the applications which are made. As to subject-matter, a chapter on heat engines might usefully be included. Not only does thermodynamics here have its primitive, and still, perhaps, its most important application, but an introductory account of the subject might help to bridge the gap between engineering and physical thermodynamics. For ease of reference in the book, it would be advantageous to have the chapter and section numbers indicated at the top of the pages.

The book as a whole is one which can be strongly recommended to students of physics and chemistry, and to all those who require a sound introductory treatise on thermodynamics.

E. C. S.

Stellar Dynamics. By W. M. SMART, M.A., D.Sc. [Pp. 434.] (Cambridge University Press. Price 30s.)

IN this work Professor Smart deals with those mathematical researches whose purpose is to bring into order the great volume of observations about the motions and distances of stars. Nearly all belong to the present century and are subsequent to the discovery of two star streams. In an introductory chapter he describes clearly the variety of objects found in the sky, and gives preliminary formulæ dealing with coordinates and magnitudes.

He begins with Eddington's "drift" function for a group of stars having haphazard velocities according to Maxwell's law in addition to a forward motion. Numerical formulæ for the number of stars in an element of angle are given, and also an expansion of the function as a Fourier series.

The Solar Motion is dealt with in several ways, including the simple method of counting given by Russell, Dagan, and Stewart. The methods of Bravais, Airy, and Kapteyn are given with the necessary limitation and caution to be used in their application. Examples are given of Boss's determination from the proper motions of the stars of his catalogue, and from the radial velocities from Lick Observatory determinations by Campbell and Moore, and by Smart and Green from the stars of Schlesinger's Catalogue. Formulæ are given to avoid numerical work by grouping the proper motions or radial velocities over large areas. It seems to the reviewer preferable to write down equations of condition for all the stars in an area of 400-600 square degrees, and form the normal equations from these means.

The beautiful researches of Eddington and Schwarzschild on the two drift and ellipsoidal distribution of velocities are given in detail. They agree satisfactorily in their determination of the vertex and apex, and give an equally good representation of the observations. The possibility of three unequal axes is dealt with, but Charlier's researches show that two of the axes are nearly equal. The next chapter is devoted to the statistical determination of the mean parallaxes of groups of stars from their peculiar proper motions or radial velocities. This statistical method is necessary to give the absolute magnitudes of Cepheid variables. It is clearly shown what caution should be shown in the application of this method.

Two following chapters are given to the variation of density of stars with distance, and to the determination of formulæ for the mean parallax of groups of stars of given magnitude and proper motion. These researches have rather lost their significance since the discovery of the rotation of the galaxy. In conjunction with spectral type they give correct indications for distances less than 500 parsecs.

A very interesting chapter follows on star-clusters. The parallax of a converging cluster is found from geometrical considerations. The distances of open clusters have been determined by Trumpler from the photographic magnitudes of stars of known spectral type compared with their absolute magnitudes. The distance of the globular clusters has been determined by Shapley from the Cepheid variables contained in them. In spite of their great distance he saw that they were associated with the Galaxy and from their one-sided distribution placed their centroid and the centre of the Galaxy at a distance of 15,000 parsecs in longitude 325° . These globular clusters are nearly of spherical form and contain thousands of stars. The distribution of the stars in them has to be inferred from their projected distribution on a photograph, and involves the solution of an integral equation. From the analogy of a spherical mass of gas,

Plummer derived a law of $B/(1+r^2)^{5/2}$. This implies a number $\frac{N}{2} x/(1+x^2)^{1/2}$

in a strip from the centre of the cluster to a distance x , and is in good accordance with observation. The dynamics of clusters is considered, and their rate of dissolution is shown to be very slow even when the rotation of the Galaxy is considered.

The next chapter on the dynamics of stellar systems opens with the frequency of collisions and proceeds to general propositions due to Eddington and Jeans.

The last two chapters are given over to the rotation of the Galaxy. After an interesting historical summary, in which contributions made by Charlier as early as 1913, Strömberg and Lindblad, we come to Oort's work in 1927. Assuming the stars to move in circular orbits under a central gravitational force, he was able to find in the radial velocities of stars at a considerable distance a differential effect indicated by a term $\bar{r} = A \sin 2(G - G_0)$. The value of G_0 agreed with the centre of the Galaxy as determined by Shapley from globular clusters. Oort confirmed his hypothesis from the proper motions and found $\bar{\mu} = A \cos 2(G - G_0) + B$. Assuming circular motion round the centre of the Galaxy, and taking $\frac{V^2}{R} = K(R)$ the attractive force to the centre,

$$A = \frac{1}{2} \frac{V}{R} - \frac{1}{2} \frac{dV}{dR}; \quad B = A - \frac{V}{R}.$$

The existence of A necessarily implies a central mass. An ellipsoidal distribution of matter would cause rotation like a solid body.

All this is clearly explained by Professor Smart with the derived dimensions and mass of the Galaxy. In the following chapter he gives Lindblad's demonstration of the ellipsoidal distribution of proper motions of stars near the Sun. To the reviewer this is entirely different from a Schwarzschild distribution of absolute velocities over the whole galaxy, with which it is connected in Chapter 12, and for which there is no observational foundation.

We have given a brief outline of the scope of Professor Smart's book, particularly of those parts admitting of comparison with observation. It was time that the researches of the present century on stellar movements were collected. Professor Smart is well fitted for the task and has produced a very useful book.

F. W. DYSON.

Demonstration Experiments in Physics. By E. SUTTON. (McGraw Hill Publishing Co.)

THE volume of works on physics which are pouring from the world's presses is so large and is increasing at such a rapid rate that it is, in general, difficult to select any one work as being novel in its outlook and indispensable to the reader. Professor Sutton's treatise does, however, possess these qualities of novelty and indispensability. There are few books on lecture-demonstration physics in the field, and none that covers so wide a range as this book does. Nearly twelve hundred experiments are listed, divided between mechanics and general physics, sound, heat, magnetism and electricity, light and atomic physics. Many of the experiments are old friends, many will be new to English readers, all are interesting and stimulating. One important and arresting feature of the book is the number of experiments which are intended as much to provoke queries as to answer them. The descriptions are necessarily brief, but are for the most part adequate, though now and again the text is so much condensed that essential points concerning the dimensions, for example, of constituent portions of the apparatus, are left for the demonstrator to investigate.

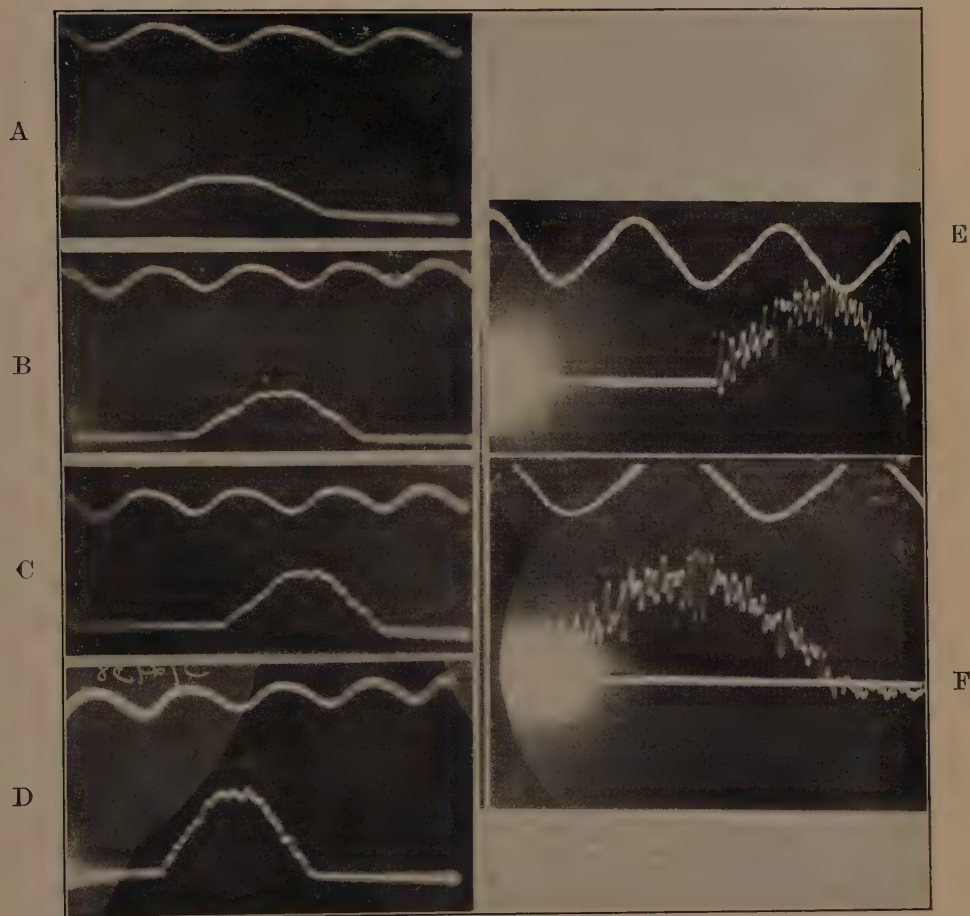
The volume will undoubtedly ease the task of the hard-pressed demonstrator seeking new experiments to illustrate fundamental physical phenomena, and should be assured of success.

A. F.

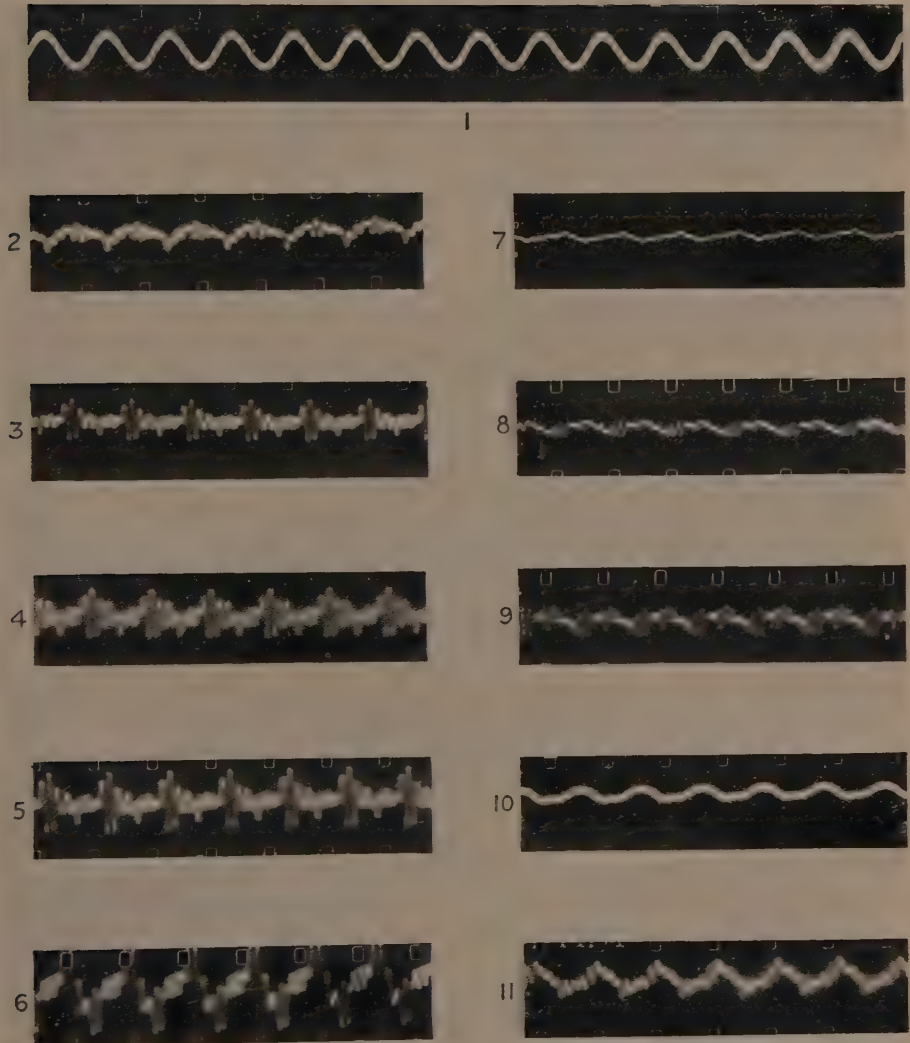
[The Editors do not hold themselves responsible for the views expressed by their correspondents.]

SOFT HAMMER.

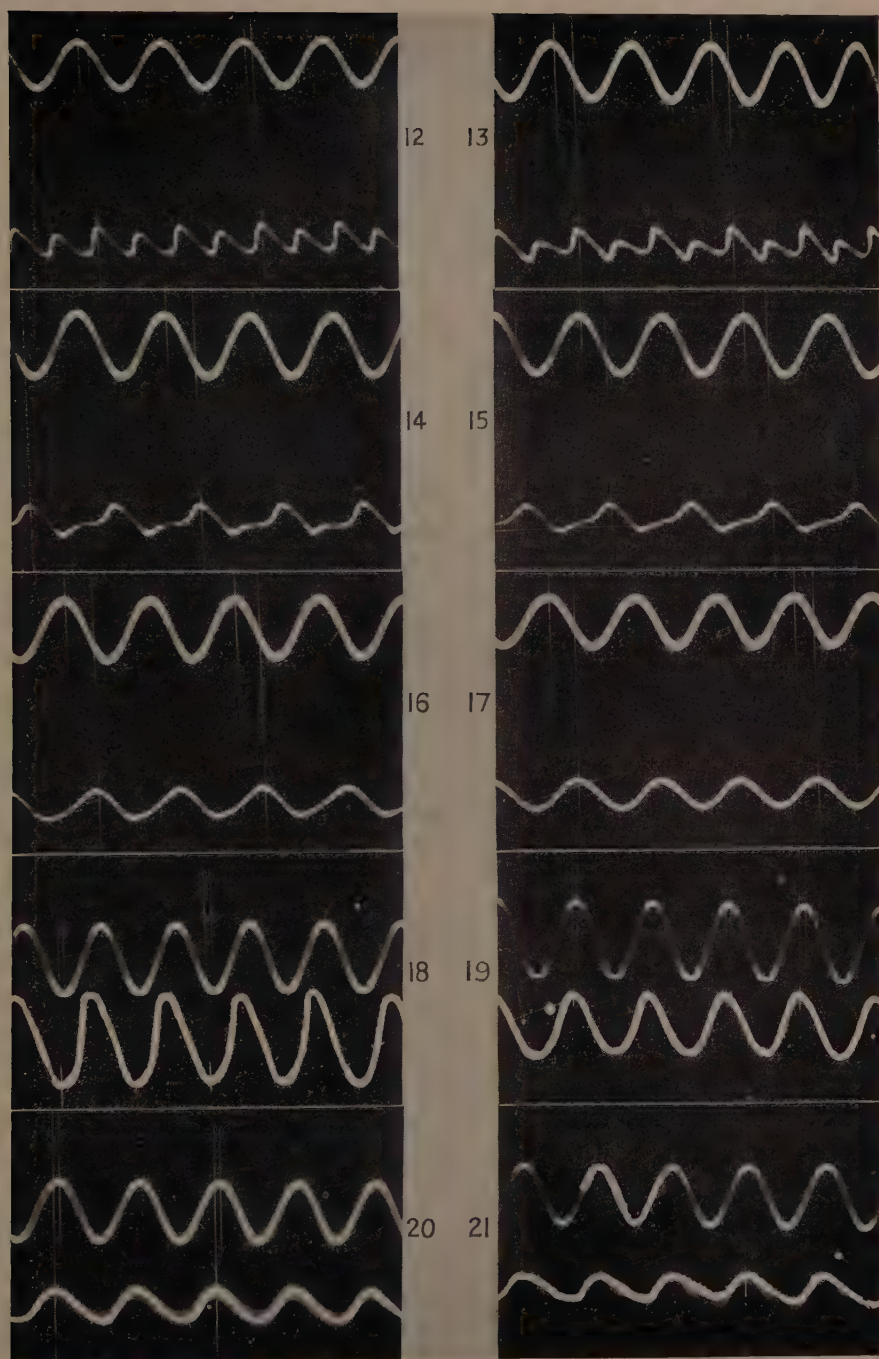
HARD HAMMER.



The lower curves in figures A, B, C, and D are oscillograms of four impacts in which the momentum changes increase from A to D, using the same soft hammer in each case. Impact commences at the left-hand foot of each hump. Abseissæ are proportional to time, ordinates to instantaneous force. In B, C, and D the distances representing one second are the same, in A the distance is greater. The lower curves in E and F are hard hammer oscillograms. The momentum change was the same in each, but the time scale is greater in F. E shows more detail at the beginning, F more at the end of impact. The six upper curves are time calibration ripples of 50 cycles per second.



Effect of Blowing Pressure on Wave Form of Reed-Pipe.



Relative Phase of Reed Vibration (*above*) and Air Vibration (*below*)
in Reed-Pipe for Various Pipe Lengths.

XXVI. *Further Note on the Propagation of Radio Waves over a Finitely Conducting Spherical Earth.*

By BALTH. VAN DER POL, D.Sc., and H. BREMMER, D.Sc.,
Natuurkundig Laboratorium der N. V. Philips' Gloeilampenfabrieken,
Eindhoven, Holland *.

[Received November 8, 1938.]

Summary.

In a former article (III.) it was pointed out that in the theory of the propagation of radio waves over a finitely conducting earth two different approximations for the many *Bessel* functions occurring in the problem can be introduced, viz., (a) the more simple *tang.* approximation (up till now used by most workers in this field), and (b) the less simple *Hankel* approximation, but which is required when the wave-length is small compared with the circumference of the earth. In III. the field on the surface of the earth due to an emitter on the ground was calculated with an exclusive use of this better *Hankel* approximation, whereas in II., the superiority of the approximation not being fully realized yet at the time, curves for the field above the ground were calculated with the aid of both approximations. The results obtained there have now been fully recalculated with the exclusive use of the *Hankel* approximation resulting in the figs. 5 and 6, which therefore should replace the former figs. 19 and 20 of II. In these new figures the curves could be given ranging from the immediate proximity of the emitter up to well into the shadow region, due to the good fit existing in the transition region between the geometric optical formula (5) and the residue formula (4 b). Here the formerly introduced "divergence factor" was numerically of great importance.

§1. *Introduction.*

IN three former papers I., II., and III. † we developed the theory of the propagation of radio waves round a spherical finitely conducting earth. Two different ways of attack were worked out:—

- (A) An extension of a series solution first given by G. N. Watson, which is best applicable in the "shadow region" of the emitter and which, further, will be called the Residue-method, and

* Communicated by the Authors.

† Phil. Mag. xxiv. p. 141 (1937); xxiv. p. 825 (1937); xxv. p. 817 (1938).

- (B) an approximate geometric-optical ray expression obtained from the rigorous solution of the problem, which was best applicable in the lit region, and which will, further, be called the geometric-optical method.

In III. we stressed the point that two different approximations to the Bessel functions occurring in the solution (A) could be made use of, viz., what we called (a) the *tang.* approximation (I. (16) and III. (10 a)), and (b) the *Hankel* approximation (I. (17), and III. (10 b)), of which, as pointed out in III., the latter is decidedly better for the practical condition $(k_1 a)^{2/3} \gg 1$, although the *tang.* approximation leads to a simpler treatment of the problem. We here use the same notations as in these former papers.

In III. (figs. 3, 4, 5, and 6) we gave a complete set of practical curves for the field on the surface of the earth ($h_2=0$) and when the emitter is also on the ground ($h_1=0$). These data were all calculated with the aid of the *Hankel* approximation only and therefore constitute a very accurate result.

As to the numerical results for an elevated emitter and (or) receiver (II., figs. 19 and 20), these were obtained partly still with the less accurate *tang.* approximation, the superiority of the *Hankel* approximation not being known yet at the time. We therefore have repeated these numerical investigations for the case of either or both the emitter and receiver at a height of $h=100$ metres and for $\sigma=10^{-13}$, $\epsilon=4$, but now with the exclusive use of the *Hankel* approximation. Moreover, we have taken the opportunity to apply our former geometric-optical solution (B) also numerically, and have obtained an excellent fit between the numerical results calculated according to the methods (A) and (B), so that we could considerably both extend and improve our former data II., figs. 19 and 20, which are therefore to be replaced by the present figures 5 and 6.

In our former investigations we could show that, contrary to the general belief, the waves bend smoothly over the horizon and have still a considerable amplitude in the "shadow region." The present extension of our former investigations fully confirms this conclusion, and waves of the order of one centimetre are beyond the horizon even stronger than we thought before. The present investigation also fully confirms our former conclusion that a full cut-off at the horizon only occurs for waves of the order of one millimetre and shorter.

§ 2. Recapitulation of Former Results.

We showed that, as a consequence of the strong absorption in the earth, the total field Π_{tot} (even behind the horizon) could be approximated to by the direct radiation $I_{-1,0}^e$ and the once reflected radiation $O_{-1,0}^e$, where

it must be remembered that both these radiations penetrate by diffraction into the "shadow region." This leads in case (A) to the series II. (77), viz. :—

$$\begin{aligned} \Pi_{\text{tot.}} &\sim I_{-1,0}^e + O_{-1,0}^e \\ &= \frac{4\pi i}{(k_1 a)^3} \sum_{s=0}^{\infty} \frac{n_s \cdot e^{in_s \pi}}{\left(\frac{\partial M_{n-1/2}}{\partial n} \right)_{n=n_s}} \cdot f_s(h_1) \cdot f_s(h_2) \cdot P_{n_s-1/2}\{\cos(\pi-\vartheta)\}, \quad (1) \end{aligned}$$

where (see II., (76)) the position of the orders n_s is determined by

$$M_n = \left[\frac{1}{x} \cdot \frac{d}{dx} \log \{x \zeta_n^{(1)}(x)\} \right]_{x=k_1 a} - \left[\frac{1}{x} \cdot \frac{d}{dx} \log \{x \zeta_n^{(2)}(x)\} \right]_{x=k_2 a} = 0. \quad (2)$$

The factors

$$\left. \begin{aligned} f_s(h_1) &= \frac{\zeta_{n_s-1/2}^{(1)}(k_1 b)}{\zeta_{n_s-1/2}^{(1)}(k_1 a)} \\ f_s(h_2) &= \frac{\zeta_{n_s-1/2}^{(1)}(k_1 r)}{\zeta_{n_s-1/2}^{(1)}(k_1 a)} \end{aligned} \right\} \dots \dots \dots (3)$$

occurring in (1) may be called the height gain factors, which can be used to extend our formula III. (20 b) (which was meant for $h_1 = h_2 = 0$) to

$$\Pi_{\text{tot}} \sim 2 \frac{e^{ik_1 a \vartheta}}{ik_1 a \vartheta} \cdot \sqrt{2\pi i} \chi \cdot \sum_{s=0}^{\infty} f_s(h_1) \cdot f_s(h_2) \cdot \frac{e^{i\tau_s \chi}}{(2\tau_s - \delta^2)}, \dots \dots (4b)$$

which expression therefore constitutes our present basis as regards the method (A).

On the other hand, the application of the geometric-optical treatment (B) (see II. (91)) leads to

$$\Pi_{\text{tot}} \sim \frac{e^{ik_1 D}}{ik_1 D} + \alpha_{11} \cdot R_{11} \cdot \frac{e^{ik_1 (R_1 + R_2)}}{ik_1 (R_1 + R_2)}, \dots \dots (5)$$

being the sum of the incident ray and the reflected ray, the amplitude of the latter being given by the product of the "divergence factor" α_{11} and the "spherical reflexion coefficient" R_{11} .

§ 3. The Height-gain Factors $f_s(h_1)$ and $f_s(h_2)$.

In II. these factors f_s were only explicitly given in the cases $\delta = 0$ and $\delta = \infty$ where

$$\delta = \frac{ik_2^2/k_1^2}{(k_1 a)^{1/3} \sqrt{k_2^2/k_1^2 - 1}}.$$

In the general case the *tang.* approximation leads to

$$f_s(h_1) \sim \frac{(-1)^s \sqrt{1-2\tau_s \delta^2}}{\delta \sqrt[4]{-2\tau_s(\chi_1^2-2\tau_s)}} \cdot \cos \left\{ \frac{\pi}{4} + \frac{1}{3}(\chi_1^2-2\tau_s)^{3/2} \right\} \\ \sim \frac{(-1)^s \sqrt{1-2\tau_s \delta^2}}{2\delta \sqrt[4]{-2\tau_s(\chi_1^2-2\tau_s)}} e^{i\pi/4 + i/3(\chi_1^2-2\tau_s)^{3/2}}, \quad \dots \quad (6a)$$

where

$$\chi_1 = (k_1 a)^{1/3} \cdot \sqrt{\frac{2h_1}{a}},$$

and

$$\chi_2 = (k_1 a)^{1/3} \cdot \sqrt{\frac{2h_2}{a}},$$

and a similar expression for $f(h_2)$. As explained above, in II. a combined use was made of the *tang.* and the, better, *Hankel* approximation, but, due to a resonance effect, instead of an improvement, a less accurate result was obtained, which caused in II. (99), (101), and (102) wrong factors to occur. In fact, $\sqrt{\chi_1}$ and $\sqrt{\chi_2}$ are to be replaced by $\sqrt[4]{\chi_1^2-2\tau_s}$, resp. $\sqrt[4]{\chi_2^2-2\tau_s}$, and in the factors $1-\frac{96}{5}\delta\tau_s^2$ the term 1 must be omitted*. Thus if we limit ourselves for the moment to the *tang.* approximation only we obtain for $h_1 \neq 0$ and $h_2 = 0$ instead of (101):

$$\Pi_{\text{tot}} \sim \frac{e^{ik_1 a \vartheta}}{ik_1 a \vartheta} \cdot \sqrt{2\pi\chi} \delta e^{-i\pi/4} \cdot \sum \frac{(-1)^s e^{i\tau_s \chi + i/3(\chi_1^2-2\tau_s)^{3/2}}}{\sqrt[4]{2\tau_s(\chi_1^2-2\tau_s)} \cdot \sqrt{1-2\delta^2\tau_s}} \quad (7a)$$

and for both h_1 and h_2 finite instead of (102)

$$\Pi_{\text{tot}} \sim 2 \frac{e^{ik_1 a \vartheta}}{ik_1 a \vartheta} \cdot \frac{\sqrt{\pi i \chi}}{4} \cdot \sum \frac{(1-2\delta^2\tau_s)}{\left(1 + \frac{\delta}{2\tau_s} - 2\delta^2\tau_s\right)} \\ \cdot \frac{e^{i\tau_s \chi + i/3(\chi_1^2-2\tau_s)^{3/2} + i/3(\chi_2^2-2\tau_s)^{3/2}}}{\sqrt{\tau_s} \sqrt[4]{(\chi_1^2-2\tau_s)(\chi_2^2-2\tau_s)}}, \quad \dots \quad (8a)$$

both being valid for χ_1^2 and $\chi_2^2 \gg 3$ and where the former first factor $2e^{ik_1 D}/(ik_1 D)$ has been replaced by the slightly more accurate expression $2e^{ik_1 a \vartheta}/(ik_1 a \vartheta)$.

But, as said before, we consider the *Hankel* approximation as the better one of the two, and with it we obtain from II. (97) for $f_s(h_1)$:

$$f_s(h_1) \sim i \sqrt{\frac{\chi_1^2-2\tau_s}{2\tau_s}} \cdot \frac{H_{1/3}^{(1)}\{1/3(\chi_1^2-2\tau_s)^{3/2}\}}{H_{1/3}^{(1)}\{1/3(-2\tau_s)^{3/2}\}} \quad \dots \quad (6b)$$

As (see II., pp. 852 and 853) the $H_{1/3}$ of the nominator of (6b) behaves differently when the argument is $>$ or $<$, i. e., when $h_{\text{metre}} >$ or $< 56\lambda^{2/3}$,

* Mr. G. Millington, in a private communication, kindly drew our attention to this discrepancy.

we approximate these two cases separately. When $h > 56\lambda^{2/3}$ and for the case of ultra short waves, where $|\delta| \ll 1$, (6 b) can further be simplified to

$$f_s(h_1) = \frac{3}{4} \sqrt{\frac{2}{\pi}} \frac{e^{-i\pi/4 + i/3(\chi_1^2 - 2\tau_s)^{3/2}}}{\delta\tau_{s,0} \sqrt[4]{\chi_1^2 - 2\tau_s} \cdot [J'_{1/3}(1/3 | 2\tau_{s,0} |^{3/2}) + J'_{-1/3}(1/3 | 2\tau_{s,0} |^{3/2})]} \quad (9b)$$

where, for $s \geq 4$, the sum $J'_{1/3} + J'_{-1/3}$ may be approximated by

$$\frac{(-1)^{s+1} \sqrt{6}}{\pi \sqrt{s + 3/4}}.$$

This formula (9 b) was derived from (6 b) by replacing $H_{1/3}$ in the nominator by the first term of its asymptotic development, whereas in the denominator $H_{1/3}$ was first expressed in $J_{1/3}$ and $J_{-1/3}$, and thereupon the first two terms of its Taylor development in δ were retained. For h near $56\lambda^{2/3}$ more terms of the asymptotic development of $H_{1/3}$ are necessary.

On the other hand, when $h < 56\lambda^{2/3}$ we better go back to the original definition (3) of $f_s(h)$ which, making use of its differential equation (I. (19)) and of the above equation (2) for the zero's n_s , together with the approximation

$$\left[\frac{d}{dx} \log \{x \zeta_n^{(2)}(x)\} \right]_{x=k_1 a} = \frac{k_2}{k_1} \cdot \frac{1}{z^{1/3} \delta},$$

yields

$$f_s(h_1) \sim 1 + \left(1 - \frac{\delta}{z^{2/3}} \right) \frac{k_1 h_1}{z^{1/3} \delta} - (1 - z^{2/3} \delta \tau_s) \frac{k_1^2 h_1^2}{z^{4/3} \delta} \dots \quad (10)$$

being a simpler form for II. (98). Moreover, a further consideration shows that (10) leads to a minimum in the field as a function of height, which, for negligible displacement current, occurs at

$$h_{\min} = \frac{\sqrt{c}}{2\pi} \cdot \sigma_{\text{e.m.u.}}^{1/2} \cdot \lambda^{3/2},$$

confirming in a simple way a result of Eckersley and Millington*.

The height-gain factors $f_s(h_1)$ and $f_s(h_2)$, appearing in (4 b) and which belong to the Residue method (A), have thus been developed in a form suitable for numerical application. In the following paragraphs we shall consider in more detail the *divergence factor* α_{11} and the *spherical reflexion coefficient* R_{11} , both belonging to the geometric-optical approximation (B).

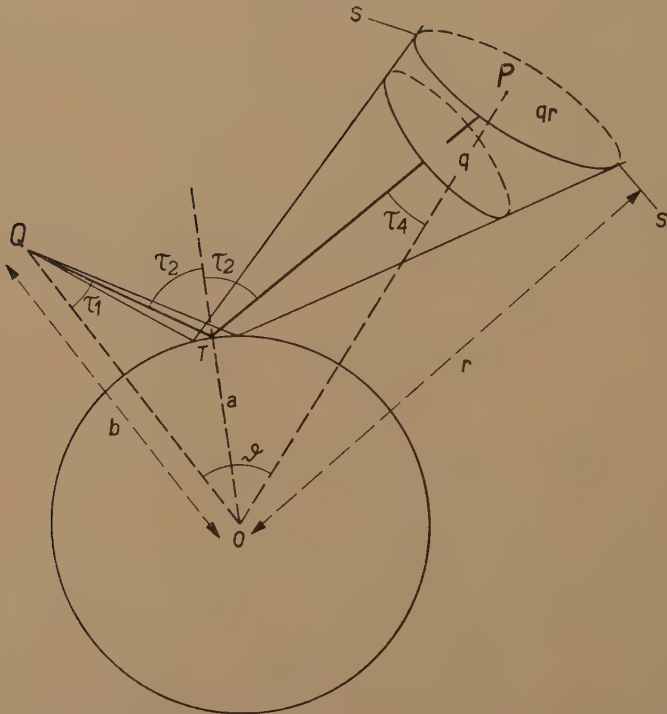
* Phil. Trans. Roy. Soc. cccxxvii. p. 297 (1938).

§ 4. The Divergence Factor α_{11} .

The divergence factor α_{11} was developed in II. (91) from the general wave solution with the aid of the saddle-point method. It was explained physically as the additional divergence of a beam of rays due to the reflexion against a curved surface. The analytical expression found for this geometrical function was

$$\alpha_{11} = \frac{(R_1 + R_2) \sqrt{k_1 a \sin \tau_2}}{k_1 b r \sqrt{\sin \vartheta \cdot \frac{\partial \vartheta}{\partial n} \cdot \cos \tau_1 \cos \tau_4}} = \frac{a(R_1 + R_2) \sqrt{\sin \tau_2 \cos \tau_2}}{\sqrt{b r \sin \vartheta (R_1 r \cos \tau_4 + R_2 b \cos \tau_1)}}, \quad \dots \dots (11)$$

Fig. 1.



Notations used in the geometric-optical formula (5).

but as this factor is of great numerical importance in determining the field before the horizon we give here another, this time purely geometrical, interpretation of this function.

Consider fig. 1, a narrow beam leaving the source Q and thereupon, after reflexion against the sphere at T, arriving at P. Let q be the actual normal cross-section of this beam at P, and let q' be the cross-section if the

reflexion had taken place against a plane surface at T, then we shall show that

$$\alpha_{11} = \sqrt{\frac{q'}{q}} \quad (12)$$

Obviously the cross-section q' at P is given by

$$q' = (R_1 + R_2)^2 \cdot \sin \tau_1 \cdot d\tau_1 d\phi,$$

where the angle τ_1 is as shown in fig. 1. The spatial angle of the beam on leaving Q is $\sin \tau_1 d\tau_1 d\phi$, where ϕ is measured at right angles to the plane of the paper. The cross-section q_r , which the beam at P cuts out from a sphere S of radius $r = OP$ and centre at O, is

$$q_r = r^2 \cdot \sin \mathfrak{S} \cdot d\mathfrak{S} d\phi.$$

Now, for fixed r , $d\mathfrak{S} = \frac{\partial \mathfrak{S}}{\partial n} \cdot dn$ (see for definition of $\frac{\partial \mathfrak{S}}{\partial n}$ II. p. 843) where

$n = k_1 b \sin \tau_1$. Further, we have $dn = k_1 b \cos \tau_1 \cdot d\tau_1$, from which it follows that

$$q_r = k_1 b r^2 \cdot \sin \mathfrak{S} \cdot \frac{\partial \mathfrak{S}}{\partial n} \cdot \cos \tau_1 \cdot d\tau_1 d\phi,$$

and also

$$q = q_r \cdot \cos \tau_1.$$

Having thus found q' and q equation (12) at once gives α_{11} , which is therefore seen to be identical with (11). Hence the divergence factor formerly derived with the aid of the saddle-point method can thus be fully explained geometrically.

As to the numerical value of α_{11} this is unity on the lit part of the surface of the sphere, but it becomes zero on the line of sight and elsewhere varies between these two limits.

§ 5. The Spherical Reflexion Coefficient R_{11} .

As shown in (5), a numerical knowledge of the spherical reflexion coefficient R_{11} is necessary in order to calculate the field in the lit part of space. The expression for R_{11} found in II. (68 a) was

$$R_{11} = \frac{-\left[\frac{1}{x} \cdot \frac{d}{dx} \log \{x \zeta_n^{(2)}(x)\}\right]_{x=k_1 a} + \left[\frac{1}{x} \cdot \frac{d}{dx} \log \{x \zeta_n^{(2)}(x)\}\right]_{x=k_2 a}}{\left[\frac{1}{x} \cdot \frac{d}{dx} \log \{x \zeta_n^{(1)}(x)\}\right]_{x=k_1 a} - \left[\frac{1}{x} \cdot \frac{d}{dx} \log \{x \zeta_n^{(2)}(x)\}\right]_{x=k_2 a}} \quad (13)$$

For a practical case, where the position of the emitter and receiver are given, first the position of the reflexion point T (fig. 1) on the surface of the earth has to be determined, so that then, also, the angle τ_2 is known. Hence in (13) we must take $n = k_1 a \sin \tau_2$, which can be written $n = a\lambda$, where $\lambda = k_1 \sin \tau_2$. Our next task is to obtain approximations of the

ζ functions in (13). As further in the $\zeta_n^{(2)}(k_2 a)$ functions the argument $k_2 a$ is never near the order $n = k_1 a \sin \tau_2$, we have, as in I., p. 152,

$$\left[\frac{d}{dx} \log \{x \zeta_n^{(2)}(x)\} \right]_{x=k_2 a} \sim \frac{\sqrt{\lambda^2 - k_2^2}}{k_2},$$

whereas for the other ζ functions, where the argument $k_1 a$ may be near the order $k_1 a \sin \tau_2$, we must use the *Hankel* approximation according to I. (22). We thus obtain, instead of (13),

$$R_{11} = \frac{k_2^2 \sqrt{\lambda^2 - k_1^2} \left[1 - \frac{k_1^2}{\lambda^2} \left\{ 1 - e^{-i\pi/6} \cdot \frac{H_{2/3}^{(2)}(u)}{H_{1/3}^{(2)}(u)} \right\} \right] - k_1^2 \sqrt{\lambda^2 - k_2^2}}{k_2^2 \sqrt{\lambda^2 - k_1^2} \left[1 - \frac{k_1^2}{\lambda^2} \left\{ 1 - e^{i\pi/6} \cdot \frac{H_{2/3}^{(1)}(u)}{H_{1/3}^{(1)}(u)} \right\} \right] + k_1^2 \sqrt{\lambda^2 - k_2^2}}, \quad (14)$$

where
$$u = \frac{a}{3\lambda^2} \cdot (\lambda^2 - k_1^2)^{3/2} \cdot e^{3/2 i \pi}.$$

Two special cases are of interest :—

- (a) $a \rightarrow \infty$, leading to the plane problem. Here (14) gives at once the well-known Fresnel reflexion coefficient in the form as already considered in I. (21).
- (b) $\tau_2 = \pi/2$ corresponding to grazing incidence. Here $\lambda = k_1$, and hence $u = 0$ leading to the special case

$$R_{11} \sim \frac{1 - e^{-2/3 i \pi} A \delta}{-1 + e^{2/3 i \pi} A \delta}, \quad (15)$$

where
$$A = \frac{6^{1/3}}{2} \cdot \frac{\Pi(2/3)}{\Pi(1/3)} = 0.919.$$

As for the ultra-short waves and average soil $|\delta| \ll 1$, we see from (15) that even for grazing incidence the spherical reflexion coefficient is very approximately equal to -1 , which is also the value of the plane reflexion coefficient.

The difference between the spherical and plane reflexion coefficient even at grazing incidence thus being very small, we may, for the conditions mentioned, everywhere in (5) replace the spherical by the ordinary plane reflexion coefficient. This, however, would not be allowed if longer waves or a reflexion against soil of a better conductivity were investigated. *E. g.*, if the conductivity were infinite, ($\delta \rightarrow \infty$), (15) would give for the value of the spherical reflexion coefficient at grazing incidence $e^{-i\pi/3}$ instead of $+1$ as for the plane case.

§ 6. Numerical Results.

Above we have given all the details for a numerical evaluation of cases of practical importance, all deductions now being exclusively based on the more accurate *Hankel* approximation.

With the aid thereof we have again with great care constructed two sets of curves giving the attenuation (as compared with the free radiation) of the waves as a function of distance, viz. :

(a) for $h_1=100$ metres, $h_2=100$ metres,

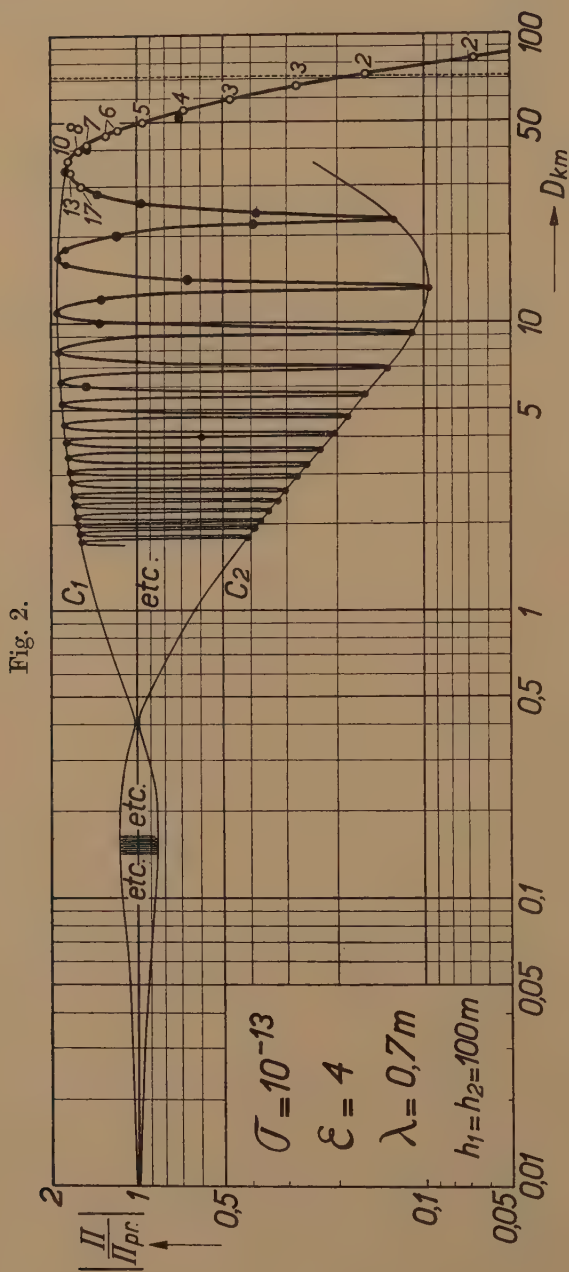
(b) for $h_1=100$ metres, $h_2=0$,

the conductivity σ and dielectric constant ϵ being taken $\sigma=10^{1-3}$ e.m.u., $\epsilon=4$ representing average soil. The range of distance covered is from 0 to well into the "shadow region"; the curves have been calculated for the wave-lengths $\lambda=7$ m., 70 cm., 7 cm., and 7 mm.

As already stated above, the parts of the curves for the lit region were calculated with the geometric-optical formula (5), the divergence factor α_{11} being duly taken into account. On the other hand, the residue formula (4 b), which is principally applicable anywhere, is most suited beyond and near the horizon, and a range of overlap with (5) exists, showing a very good numerical agreement. Thus (4 b) and (5) together enable us to draw uninterrupted field curves for distances ranging from close to the emitter up to far into the shadow region.

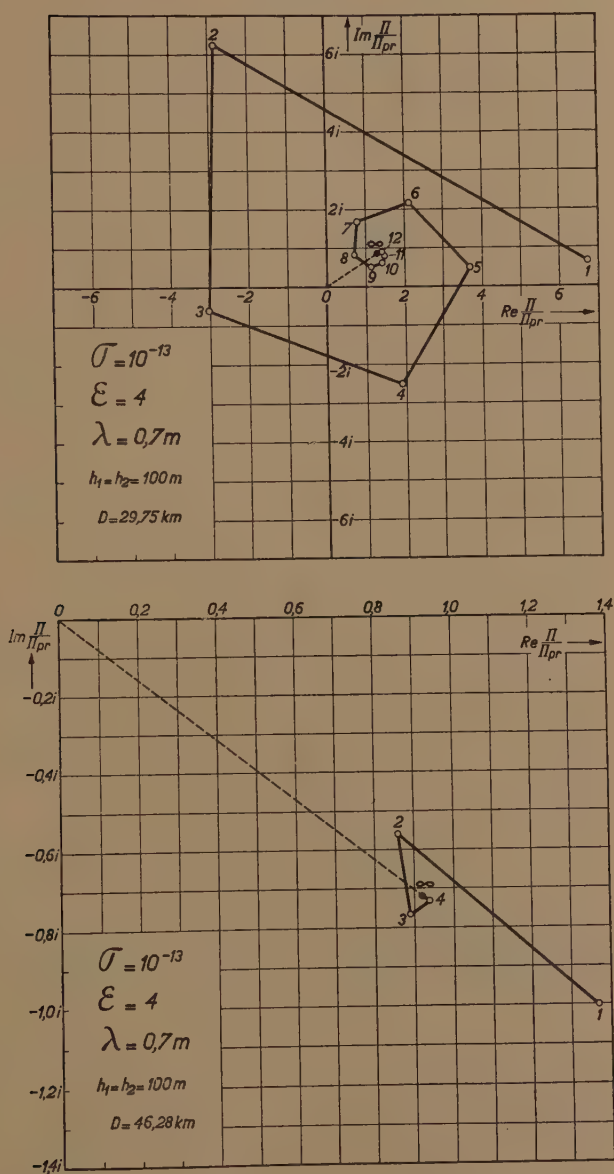
Before, however, describing the final curves we give three additional figures 2, 3, and 4, which show some detailed results of the calculation. Consider first fig. 2, which has been drawn for $\sigma=10^{-13}$, $\epsilon=4$, $\lambda=0.7$ metre, $h_1=h_2=100$ metres. In this figure, therefore, the emitter is at a height of 100 metres, whereas the receiver is thought to be all the time at the same height, the linear distance D being varied. Up to the shadow limit maxima and minima occur due to the interference of the direct and reflected radiation as represented by the first and second term of (5). Close to the emitter the ratio Π/Π_{pr} is obviously unity, but soon we reach the distance where it can be said that the interference maxima lie on the curve C_1 given by $|\Pi/\Pi_{pr}| = 1 + \alpha_{11} |R_{11}|$ and the minima on the curve C_2 , viz. $|\Pi/\Pi_{pr}| = 1 - \alpha_{11} |R_{11}|$. It is seen from fig. 2 that the distance $2\alpha_{11} |R_{11}|$ between C_1 and C_2 shows a minimum at $D=0.4$ km., this being due to the Brewster minimum in the reflexion. It may be remarked that beyond this minimum the effect of the divergence factor α_{11} becomes more and more of importance; if we had wrongly considered the reflexion to take place against a flat earth, the interference minimum at $D=13$ km., *e. g.*, would have become 0.066 instead of 0.096.

When we approach the shadow limit and penetrate into it the residue series formula (4 b) must be used. Therefore in fig. 2 all the points calculated from the geometric-optical formula (5) are shown as black circles, whereas those obtained from the residue series as open circles. The numbers placed near these open circles represent the number of terms of the residue series which were necessary in order to obtain an accuracy of one per cent. The figure shows very clearly how this number



An example ($\lambda = 70$ cm.) showing the interference phenomena in the lit region and the transition from the geometric-optical results (5) shown as • to those obtained from the residue series (4 b) shown as C.

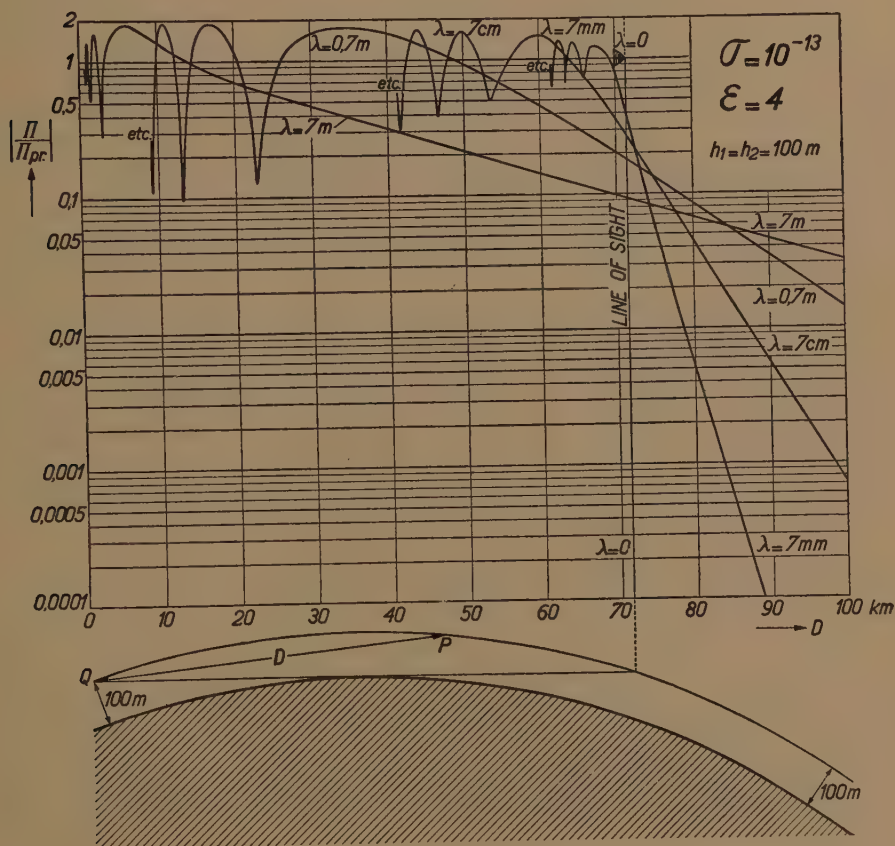
Fig. 3.



Two examples showing the evaluation of the residue series as a vectorial addition of its successive terms. The top drawing refers to a point between the last interference maximum and minimum, whereas the lower figure refers to a point between the last maximum and the shadow limit.

and the vector drawn from the origin O to any point on the spiral represents, both in amplitude and phase, the ratio of the actual field at the corresponding distance to the field which would obtain there in the absence of the diffracting earth. If we follow the spiral from the centre $(1+oi)$, which corresponds to the receiver near the emitter, it first spirals out and

Fig. 5.

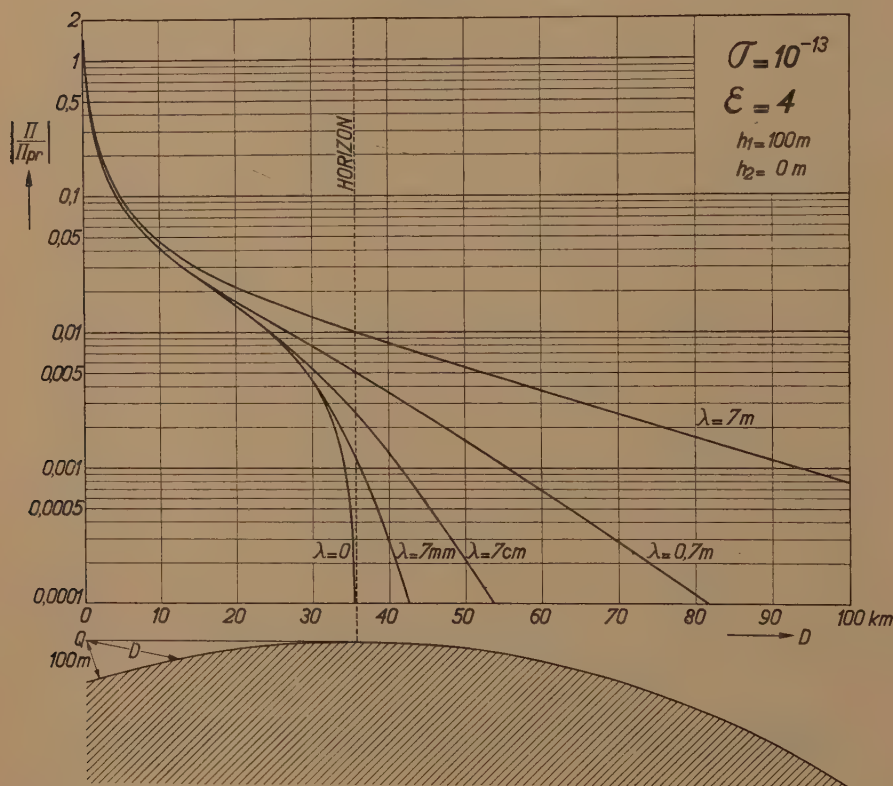


The attenuation factor for short waves when both emitter and receiver are at a height of 100 m. and for propagation over average soil.

then shrinks again to a minimum value at (B) corresponding to the Brewster minimum, as fully explained with fig. 2; next it expands again (the radius corresponding all the time to $\alpha_{11} | R_{11} |$) till it bends over at the point S, where the shadow limit is reached, finally to approach the origin O again in a spiral way, but with inversed rotation, the details of this latter part being given in the left top drawing.

Having thus explained some details of our general numerical procedure we give in figs. 5 and 6 the final results. Both figures have again been calculated for $\sigma=10^{-13}$, $\epsilon=4$, $\lambda=7$ m., 0.7 m., 7 cm., and 7 mm. Fig. 5 relates to the case where both emitter and receiver are at 100 metres height ($h_1=h_2=100$ m.), whereas in fig. 6 the emitter is supposed at $h_1=100$ metres, the receiver being on the ground ($h_2=0$), where it may be remarked that, due to the principle of reciprocity, the last figure also

Fig. 6.



The attenuation factor for short waves of the field at the surface of the earth when the emitter is at a height of 100 m. and for propagation over average soil.

refers to the case where the emitter and receiver are interchanged. Thus the present more accurate fig. 5 has to replace fig. 20 of II. and the present fig. 6, fig. 19 of II. Fig. 5 clearly shows again that a wave $\lambda=7$ metres penetrates well into the optical shadow region, no discontinuity whatever occurring at the shadow limit. As found before, only waves of a length

of the order of one centimetre or shorter exhibit a rapid cut-off when we pass into this shadow region.

The general behaviour of the curves of fig. 5 can be described as follows:—The final decrease for short waves starts already at the last interference maxima (the one nearest the shadow limit), the slope of the curve varying further only very little when we enter into the shadow region. It should be observed that this last maximum occurs the nearer the shadow limit the shorter the wave-length, it lying on this limit when $\lambda \rightarrow 0$, for which case an infinite sharp cut-off occurs.

Turning now to fig. 6, this has been drawn for the same constants of the earth and wave-lengths, only $h_1=100$ m. and $h_2=0$. Here obviously no interference maxima and minima occur, the curves being of a fully monotonic character with a complete absence again of any discontinuity at the horizon. Before the horizon the geometric-optical formula (5) was again used up to the transition region, after which the residue series (4 b) was applied. Moreover, in this case, where the receiver is on the ground, we have for all distances two simplifications in the general formula (5), viz., $\alpha_{11}=1$, and $D=R_1+R_2$.

It is further seen that before the horizon the curves asymptotically approach the limiting curve for $\lambda \rightarrow 0$, which latter case could be very simply treated with the aid of (5), because then the conduction current is negligible compared with the dielectric displacement current ($k_2^2 \rightarrow \epsilon \omega^2/c^2$), leading quite simply to

$$\frac{\Pi}{\Pi_{pr}} = \frac{2\epsilon \cos \tau_2}{\epsilon \cos \tau_2 + \sqrt{\epsilon - \sin^2 \tau_2}},$$

where τ_2 is the angle of incidence of the direct ray at the receiving point.

In conclusion, the feeling may be expressed that, as far as we are aware, the new curves given here, and which are again based on the hypothesis of a homogeneous atmosphere, constitute the most accurate and detailed theoretical results so far published because they were derived with one single approximation to the Bessel functions, viz., what we have called the *Hankel* approximation.

Eindhoven.

31st October, 1938.

XXVII. *The Thermal Stability of a Cylindrical Stratified Dielectric* *.

By S. WHITEHEAD, M.A., Ph.D. †

[Received September 3, 1938.]

SUMMARY.

The solution is given for the limit of thermal stability of a dielectric, comprised of layers with intervening equipotential surfaces, made up into a cylindrical form. A comparison is made with the same type of dielectric when made up into a flat form. Certain deductions are made as regards the particular application to condensers.

General.

IT is well known that when, as is usual, the energy losses in a dielectric increase with temperature, then, provided a sufficiently high electric stress is applied, the temperature of the whole or a part of the dielectric will rise to such a value that the insulating properties are lost and breakdown ensues ‡. If, further, the temperature coefficient of the energy losses increases with temperature, then, at a certain electric stress, the temperature becomes unstable, even though at a slightly lower stress the temperature is well within the temperature limits of the dielectric.

The solution of this problem has been worked out for a uniform disk or cylinder of dielectric. A case of some importance is when the dielectric is formed from thin layers with conducting surfaces. With plane parallel electrodes this feature does not alter the solution (unless the conducting surfaces are of appreciable thickness compared with the dielectric); nor is the solution altered with a cylindrical dielectric exposed to a continuous unidirectional ‡ potential difference, except when the dielectric is applied as a spiral roll. But in the latter case, and with alternating voltages applied to a cylindrical dielectric, the stratification in the manner generally applied causes the field to be uniform across

* Based on report Ref. L/T90, "Thermal Stability of Condensers," of the British Electrical & Allied Industries Research Association.

† Communicated by the Author.

‡ See, for example, 'Dielectric Phenomena,' vol. iii. "Breakdown of Solid Dielectrics," by S. Whitehead (Ernest Benn, 1932), p. 226, *et seq.*

‡ Thermal instability rarely occurs with d. c., so that this case is not of practical importance.

the cross-section, and therefore modifies the solution. It is the aim of this note to give the appropriate solution which has not hitherto been shown completely.

The equation for thermal equilibrium is

$$\frac{1}{r} \frac{d}{dr} \left(r \frac{d\theta}{dr} \right) = \sigma K E^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where

θ = temperature rise above ambient at radius r ,

σ = conductivity or energy losses in watts/cm.³ per volt/cm.,

K = thermal resistivity of dielectric in thermal ohms,

E = field strength in volt/cm.

A frequent type of variation of σ with temperature, and one which is usually valid over a small range, is given by

$$\sigma = \sigma_0 e^{\gamma\theta}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Let it be supposed that no heat is lost from the internal surface of the cylinder, then, if

$$\log(r/b) = x, \quad 2x + \gamma\theta = y, \quad \lambda_0 = \sigma_0 K E^2,$$

and b = external radius of cylinder,

$$\frac{dy}{dx} = -2\sqrt{(\alpha^2 - \frac{1}{2}\lambda_0\gamma b^2 e^{\gamma\theta})}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.1)$$

$$\alpha^2 = 1 + \frac{1}{2}\lambda_0\gamma a^2 e^{\gamma\theta_m}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.2)$$

where θ_m is the temperature of the internal surface, which has a radius a .

By integration and substitution for x we now obtain

$$\sqrt{(\frac{1}{2}\lambda_0\gamma r e^{\gamma\theta})} = 2\alpha \left\{ \frac{1}{m} \left(\frac{r}{b} \right)^\alpha + m \left(\frac{b}{r} \right)^\alpha \right\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.1)$$

where

$$m = \left(\frac{\alpha+1}{\alpha-1} \right)^{\frac{1}{2}} \left(\frac{a}{b} \right)^\alpha. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.2)$$

If the thermal dissipation from the external surface follows Newton's law, then

$$(\alpha^2 - \frac{1}{2}\lambda_0\gamma b^2 e^{\gamma\theta_1}) = \frac{1}{2}\beta\gamma\theta_1 - 1, \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.1)$$

where θ_1 is the temperature rise at $r=b$ and $\beta = Kb/P_e$, where P_e is the external thermal resistance, that is β corresponds to the ratio of the internal to the external thermal resistance. Further, from (4.1) we have

$$\sqrt{(\frac{1}{2}\lambda_0\gamma b^2 e^{\gamma\theta_1})} = 2\alpha m / (m^2 + 1). \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.2)$$

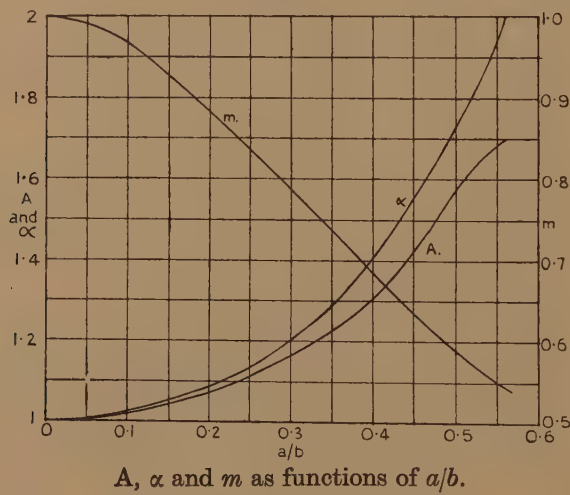
Equations (5) permit the determination of α and θ_1 , together with (4.2), whence all the other quantities can be found. The condition that thermal equilibrium is impossible is, therefore, the condition that

(5) should have no real roots. Accordingly, the limit of thermal stability is given by the condition that the Jacobian of equations (5) should vanish. This latter condition can be expressed as

$$1 = \left\{ \left(\frac{2m}{1+m^2} \right)^2 + \frac{\beta}{\alpha} \left(\frac{1-m^2}{1+m^2} \right) \right\} \left\{ 1 - \frac{1-m^2}{1+m^2} \left(\frac{\alpha}{\alpha^2-1} - \log m + \frac{1}{2} \log \frac{\alpha+1}{\alpha-1} \right) \right\} \quad (6)$$

The equations (5) and (6) determine α , θ_1 , and λ_0 , whence E may also be found. The general explicit solution of these equations is somewhat laborious, so some particular cases will be examined which actually cover the main conditions of practice.

Fig. 1.



External Surface at Constant Temperature.

For this case $\theta_1=0$ and $P_e=0$, that is $\beta \rightarrow \infty$, equations (5) and (6) reduce to

$$\sqrt{(\frac{1}{2}\lambda\gamma b^2)} = 2\alpha m / (m^2 + 1), \quad (7.1)$$

$$\frac{1+m^2}{1-m^2} + \log m = \frac{\alpha}{\alpha^2-1} + \frac{1}{2} \log \frac{\alpha+1}{\alpha-1}, \quad (7.2)$$

whence
$$E = \frac{A}{b} \sqrt{\frac{2}{\sigma_0 k \gamma}}, \quad (7.3)$$

where $A = 2\alpha m / (m^2 + 1)$.

The variation of A , α and m with a/b is given in fig. 1. When a/b is

small, α and m are unity and A is $\sqrt{2}$, thus

$$E = \sqrt{(2/\sigma_0 k \gamma b^2)} \quad \dots \dots \dots (7.31)$$

Final conductance at breakdown

Initial conductance

$$\begin{aligned} &= -\frac{2b(d\theta/dr)r=b}{K(b^2-\alpha^2)\sigma_0 E^2} \\ &= \frac{2\{\sqrt{(A^2-\alpha^2)}+1\}}{A^2\{1-(\alpha/b)^2\}} \\ &= 2 \text{ when } \alpha/b \text{ is small.} \end{aligned}$$

The ratio increases with a/b , but only slightly, being only 2.28 for $\alpha=b$.

The maximum temperature of any portion of the dielectric is given by θ_m and

$$\theta_m = \frac{1}{\gamma} \log \left[\frac{\alpha^2-1}{\alpha^2 m^2 \alpha} \left(\frac{1+m^2}{2m} \right) \left(\frac{\alpha+1}{\alpha-1} \right)^{1/\alpha} \right] \quad \dots \dots \dots (7.5)$$

When a/b is small this becomes

$$\theta_m = 2 \log 2/\gamma = 1.386/\gamma. \quad \dots \dots \dots (7.51)$$

It is to be noted that this subsection also applies to the limiting case of very large condensers irrespective of other conditions. The principal equations of the subsection have previously been given by Dreyfus * in connexion with condenser-type bushings.

Internal Radius Small Compared with External Radius.

Equation (6) now gives

$$\beta \frac{1-m^2}{1+m^2} + \left(\frac{2m}{1+m^2} \right)^2 = 0.$$

That is

$$\left. \begin{aligned} \left(\frac{2m}{1+m^2} \right)^2 &= \beta \left[\sqrt{\left(\frac{\beta^2}{4} + 1 \right)} - \frac{\beta}{2} \right], \\ \text{and} \quad \frac{1-m^2}{1+m^2} &= \frac{\beta}{2} - \sqrt{\left(\frac{\beta^2}{4} + 1 \right)}, \end{aligned} \right\} \quad \dots \dots \dots (8.1)$$

together with $\alpha=1$.

Whence

$$\begin{aligned} E &= \frac{1}{b} \sqrt{\frac{2\beta}{\sigma_0 K \gamma}} \left\{ \sqrt{\left(\frac{\beta^2}{4} + 1 \right)} - \frac{\beta}{2} \right\}^{1/2} \\ &\quad \times e^{-\frac{1}{\beta} \left[1 - \sqrt{\left(1 + \frac{\beta^2}{4} \right)} + \frac{\beta}{2} \right]} = \frac{B}{b} \sqrt{\frac{2\beta}{\sigma_0 K \gamma}} \quad \dots \dots (8.2) \end{aligned}$$

* Dreyfus, S.E.V. Bull. pp. 7 and 15 (1924).

The parameter B is shown in fig. 2 as a function of β . When β is large, B tends to $\sqrt{1/\beta}$, and

$$E = \frac{1}{b} \sqrt{\frac{2}{\sigma_0 K \gamma}} \dots \dots \dots (8.21)$$

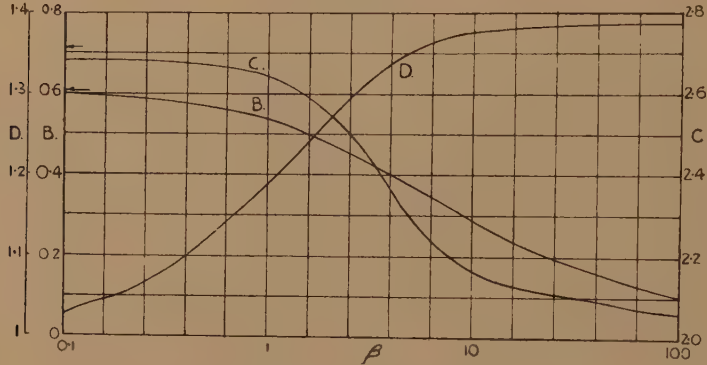
as before.

When β is small, B tends to $\sqrt{1/e}$, and

$$E = \frac{1}{b} \sqrt{\frac{2\beta}{\sigma_0 K \gamma e}} = \sqrt{\frac{2}{\sigma_0 b P_e \gamma e}} \dots \dots \dots (8.22)$$

This equation applies when the external thermal resistance is high compared with that of the dielectric, and may be deduced directly by

Fig. 2.



B, C, and D as functions of β .

a simple consideration of thermal balance. If a is appreciable, b may be replaced by

$$b[1-(a/b)^2].$$

$$\frac{\text{Final conductance}}{\text{Initial conductance}} = \frac{2\beta\theta_1}{\lambda_0(b^2-a^2)}$$

$$\times \frac{2}{\beta} \left[\frac{1}{\sqrt{(\frac{\beta^2}{4} + 1)} - \frac{\beta}{2}} - 1 \right] e^{2\{1 - \sqrt{(\frac{\beta^2}{4} + 1)} + \frac{\beta}{2}\}/\beta}$$

=C. (Given in fig. 2.) $\dots \dots \dots (8.3)$

When β is large the ratio tends to (2), as follows from the previous section. When β is small and the thermal resistance is outside the dielectric the ratio of losses tends to e or 2.718.

The maximum temperature θ_m is given by equation (7.5) as before, which becomes for the present case

$$\begin{aligned}\theta_m = & \frac{1}{\gamma} \left[\frac{2}{\beta} \left\{ 1 - \sqrt{\left(\frac{\beta^2}{4} + 1 \right)} + \frac{\beta}{2} \right\} \right. \\ & \left. + 2 \log \left\{ 2 / \left[1 + \sqrt{\left(\frac{\beta^2}{4} + 1 \right)} - \frac{\beta}{2} \right] \right\} \right] \\ = & D/\gamma. \quad \dots \dots \dots (8.4)\end{aligned}$$

The parameter D is also shown in fig. 2. It varies between $2 \log 2$ as previously, when β is large, and unity when β is small and when the thermal resistance is outside the specimen.

Comparison of Flat and Cylindrical Dielectrics.

Considering for simplicity the case when a is small, then

$$E(\text{flat})/E(\text{cylindrical}) = \eta e^{-\mu/2}/B\sqrt{\beta}, \quad \dots \dots \dots (9)$$

where

$$(1 + \beta^{-1})(\tanh \eta/\eta) + \operatorname{sech}^2 \eta = \eta^{-2}, \quad \dots \dots \dots (9.1)$$

and

$$\mu = (2/\beta)\eta \tanh \eta + \log \cosh^2 \eta. \quad \dots \dots \dots (9.2)$$

The quantities η and μ may be found elsewhere. The value of (9) varies only between 0.707 (β small) and 0.665 (β large).

If the dielectric forms a condenser, account must be taken of the greater capacity possible, with a flat condenser, for a given thickness. For a given capacity (9) becomes

$$E(\text{flat})/E(\text{cylindrical}) = b\eta e^{-\mu/2}/\delta\beta\sqrt{\beta}, \quad \dots \dots \dots (9.3)$$

while β is replaced by $\delta\beta/b$ in (9.1) and (9.2), where 2δ is the thickness of the flat condenser. As an example, let β be about 0.5, then (9.3) will not greatly exceed unity until the breadth of the flat condenser exceeds about three times the thickness :

$$\frac{\text{loss ratio (flat)}}{\text{loss ratio (cyl.)}} = \frac{\tanh \eta \cosh^2 \eta e^{2\eta \tanh \eta/\beta}}{C}. \quad \dots \dots \dots (10)$$

The ratio in equation (10) varies from unity when β is small to 1.135 when β is large ; while also

$$\frac{\theta_m(\text{flat})}{\theta_m(\text{cyl.})} = \frac{\mu}{D}. \quad \dots \dots \dots (11)$$

This ratio varies from unity when β is small to 0.856 when β is large.

It may therefore be concluded in practice that a consideration of cylindrical condensers may be taken as also applying substantially to condensers of a flattened or oval shape.

Magnitude of Highest Stable Stress.

Consider the case of a cylindrical condenser where b/a is large, then, as in equation (8.2),

$$E = (B/b) \sqrt{(2\beta/\sigma_0 K \gamma)}.$$

For A.C. of 50 c.p.s.

$$\sigma_0 = 0.2778 \times 10^{-10} \epsilon_0 \tan \delta_0,$$

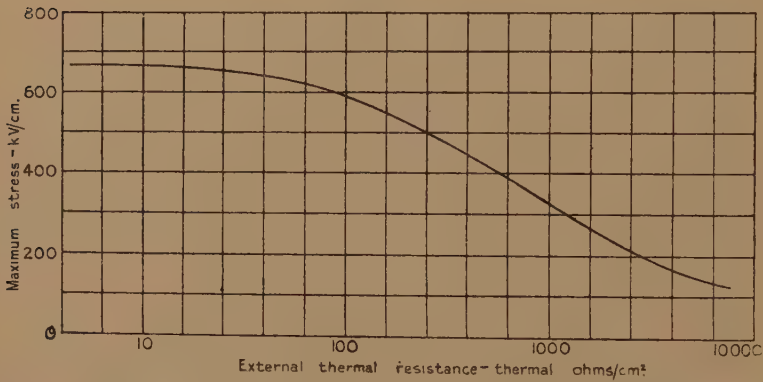
where ϵ_0 = permittivity and δ_0 = loss angle at ambient temperature. Thus

$$E = (B/b) \sqrt{(7.2 \beta / K \gamma \epsilon_0 \tan \delta_0)} \times 10^5 \text{ volts/cm.} \quad \dots (12)$$

As illustration, representative values, at room-temperature, for impregnated paper dielectric may be taken as

$$\epsilon_0 = 4.5; \tan \delta_0 = 0.002; \gamma = 0.024; K = 750,$$

Fig. 3.



Effect of external thermal resistance on stable stress.

whence—

$$E = \frac{B\sqrt{\beta}}{b} \times 6.67 \times 10^5 \text{ volts/cm.} \quad \dots (12.1)$$

Consider first the variation of E with the external thermal resistance for a given size, that is, take a condenser of 1 cm. external diameter so that β is $750/P_e$ or P_e is $750/\beta$ thermal ohms per cm.² Fig. 3 shows the variation of E with P_e in these circumstances. As the external thermal resistance is reduced E reaches a maximum of 667 kV/cm.

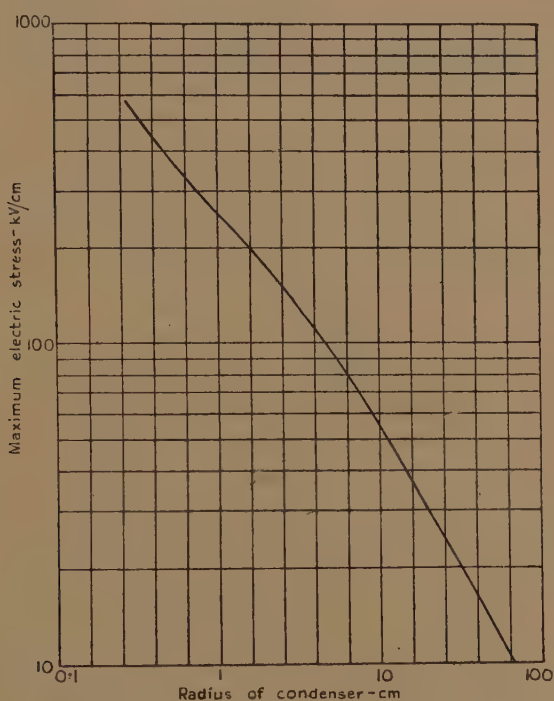
In examining the effect of diameter, some external thermal resistance must be assumed. If $P_e = 0$ then B is $\beta^{-1/2}$, and

$$E = 667/b \text{ kV/cm.,} \quad \dots (12.2)$$

that is, the stress is inversely proportional to the diameter. Again, the condenser may be assumed to be cooled by free convection and radiation. A representative value for P_e is then $1/0.0006$ and β is $0.45 b$. In these circumstances fig. 4 shows the variation of maximum electric stress with diameter, indicating the nature of the decrease of the former with increase of the latter.

However, the greater the diameter the greater the volt-ampere rating, and it is of interest to take this into account. For simplicity, attention will be confined to very large condenser units, or to units the exterior

Fig. 4.



Effect of diameter on air cooled condenser.

of which is cooled to the ambient temperature. This is the most favourable case. If t is the thickness of the dielectric (neglecting the electrode thickness) then the capacity is

$$(\epsilon b^2/4t^2) \times 1.1 \times 10^{-12} \text{ farads per cm. length.}$$

The maximum volt-amperes per cm. length passed by the condenser are ($\omega = 2\pi$ frequency)

$$\begin{aligned} VA &= (Et)^2 (\epsilon b^2/4t^2) \times 1.1 \times 10^{-12} \omega \\ &= E^2 b^2 \epsilon 0.2778 \times 10^{-12} \omega. \end{aligned}$$

But, from equation (12), remembering B is $\beta^{-1/2}$ for this case,

$$E^2 b^2 = 200\pi / 0.2778 \times 10^{-10} K \gamma \omega \epsilon_0 \tan \delta_0.$$

Thus, since ϵ is substantially equal to ϵ_0 ,

$$VA = 2\pi / K \gamma \tan \delta_0 \text{ volt-amperes/cm. length.}$$

In other words, the upper limit of the permissible volt-amperes per unit is a constant of the material only. Inserting the previous representative values for K, γ , and σ ,

$$VA = 174 \text{ volt-amperes/cm. length.} \quad . \quad . \quad . \quad (13.1)$$

It may be noted that if, on the contrary, the external thermal resistance is high, E is proportional to $b^{-1/2}$, so that the volt-ampere rating will increase proportionately with the square-root of the diameter.

Limitations of Temperature and Stress.

The maximum temperature must not exceed any critical disintegration or decomposition temperature of the dielectric. Let such a temperature be θ_c and the ambient temperature be θ_0 . Then the condition of remaining within the temperature range of the dielectric is

$$\theta_c > \theta_0 + (D/\gamma) \text{ for cylindrical form,} \quad . \quad . \quad . \quad (14.1)$$

and

$$\theta_c > \theta_0 + (\mu/\gamma) \text{ for flat form.} \quad . \quad . \quad . \quad (14.2)$$

This condition will usually be fulfilled except at high ambient temperatures.

Above a certain stress the majority of dielectrics exhibit an increase of conductivity or energy losses with stress. There may then often be applied a simple first approximation, namely, that if the effect of stress on the conductivity can be represented by a multiplying factor $f(E)$, then σ_0 can be replaced by $\sigma_0 f(E)$ in the final results already given.

In the experimental determination of thermal instability, the development of an unstable increase in the properties of the dielectric for stresses not greatly exceeding the limiting stress is often very slow. From equations (8.3) and (10) it can be seen that if, after application of voltage, the energy losses increase more than two or three times, depending on the conditions of test, the condenser is unstable. If, when the energy losses have so risen above the value calculated, the stress is then reduced, the losses will fall towards a value less than two or three times the initial value, if this new stress is within the stable limit. By this means an estimate of the maximum stable voltage can be obtained without undue delay. Further, if, in a type test, the increase of losses with time is limited to a small fraction, it can be assumed that thermal instability of the type discussed will not occur in the conditions specified.

Conclusions.

(a) The conditions for thermal stability, and in addition, the limiting electric stress in the dielectric thereby arising, can be conveniently computed for a stratified dielectric by the methods set out therein.

(b) Under the most favourable thermal conditions, the maximum volt-ampere (per unit length) rating of a condenser with respect to thermal stability depends only on the thermal resistivity and power factor of the dielectric and the variation of the latter with temperature.

(c) A rapid method of test for thermal stability is indicated, and methods are shown by which the effect may be evaluated of limitations due to the temperature of decomposition or disintegration of the dielectric and limitations due to the variation of conductivity with electric stress.

XXVIII. *Some Problems of Finite Strain.*—I.

By B. R. SETH, M.A., D.Sc.*

[Received June 27, 1938.]

IN two recent papers † the theory of finite strain in elastic problems has been developed on the hypothesis that the second order terms in the components of strain may not be neglected. Like the body-stress equations, these components have been referred to the actual position of a point P of the material in the strained condition, and not to the position of a point considered before strain, as has been invariably done by previous investigators. No account has been taken of the generalized Hooke's Law. The only assumption which has been made with regard to the stress-strain relations is that in isotropic elastic bodies they are in the simplest tensor form that we can take, viz.,

$$\bar{x}\bar{x} = \lambda(S_x + S_y + S_z) + 2\mu S_x,$$

$$\bar{y}\bar{z} = \mu\sigma_{yz}, \text{ etc.},$$

in the usual notation. As may be expected, the corresponding tension-stretch curve of the ordinary theory is now not a straight line, but it is not unlike that which is actually found in some materials ‡.

This method has already been applied to many problems, including those where the ordinary theory cannot give any satisfactory result. To give an example §, we may mention the case of a thick cylindrical tube turned inside out under no surface tractions. The object of the present paper is to extend the theory to the following cases :—

(a) A cylinder under radial body force, with special reference to a rotating shaft.

(b) A spherical shell under radial body force, with special reference to a gravitating shell.

* Communicated by the Author.

† Seth, Phil. Trans. Roy. Soc. A, cexxxiv. pp. 231–264 (1935); Shephard and Seth, Proc. Roy. Soc. A, clvi. pp. 171–192 (1936).

‡ *Loc. cit.* Phil. Trans. Roy. Soc. p. 237.

§ *Loc. cit.* Proc. Roy. Soc. pp. 186–190.

(a) *Cylinder under Radial Body Force.*

Let μR be the radial body force acting on the cylinder, μ being the rigidity. Since the cylinder is strained symmetrically the displacement is purely radial. Hence we can put

$$u=x(1-\beta), v=y(1-\beta), w=0, \quad . \quad . \quad . \quad (1)$$

where β is a function of $r=(x^2+y^2)^{\frac{1}{2}}$ only.

Here we have assumed that the longitudinal tension is adjusted so that the length is maintained constant. If we make the resultant longitudinal tension vanish, we have only to put $w=\alpha z$, where α is a constant.

The strain components which do not vanish are given by

$$\begin{aligned} S_x &= \frac{\partial u}{\partial x} - \frac{1}{2} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 \right] \\ &= \frac{1}{2} (1-\beta^2) - \frac{1}{2} x^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right), \quad . \quad . \quad . \quad (2.1) \end{aligned}$$

$$S_y = \frac{1}{2} (1-\beta^2) - \frac{1}{2} y^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right), \quad . \quad . \quad . \quad (2.2)$$

$$\begin{aligned} \sigma_{xy} &= \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} - \left[\frac{\partial u}{\partial x} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} \right] \\ &= -xy \left(\beta'^2 + \frac{2\beta\beta'}{r} \right), \quad . \quad . \quad . \quad (2.3) \end{aligned}$$

where

$$\beta' = d\beta/dr.$$

The stress components which do not vanish are given by

$$\widehat{x x} = \lambda \delta + \mu \left[1 - \beta^2 - x^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right) \right], \quad . \quad . \quad . \quad (3.1)$$

$$\widehat{y y} = \lambda \delta + \mu \left[1 - \beta^2 - y^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right) \right], \quad . \quad . \quad . \quad (3.2)$$

$$\widehat{z z} = \lambda \delta, \quad . \quad . \quad . \quad (3.3)$$

$$\widehat{x y} = -\mu x y \left(\beta'^2 + \frac{2\beta\beta'}{r} \right). \quad . \quad . \quad . \quad (3.4)$$

In polar coordinates these become

$$\widehat{r r} = \lambda \delta + \mu [1 - (r\beta' + \beta)^2], \quad . \quad . \quad . \quad (3.5)$$

$$\widehat{\theta \theta} = \lambda \delta + \mu [1 - \beta^2], \quad . \quad . \quad . \quad (3.6)$$

$$\widehat{r \theta} = 0, \quad . \quad . \quad . \quad (3.7)$$

where

$$\delta = S_x + S_y + S_z = 1 - \beta^2 - \frac{1}{2} (r^2 \beta'^2 + 2r\beta\beta').$$

In order that these stresses may satisfy body-stress equations of the type

$$\frac{\partial \widehat{xx}}{\partial x} + \frac{\partial \widehat{xy}}{\partial y} + \mu \rho R \frac{x}{r} = 0,$$

we find that β should satisfy the differential equation

$$\beta^2 + (r\beta' + \beta)^2 + c \int r\beta'^2 dr - c \int \rho R dr = K, \quad (4)$$

where

$$c = \frac{2\mu}{\lambda + 2\mu} = \frac{1 - 2\eta}{1 - \eta},$$

and K is a constant of integration.

To simplify the results we shall assume that the cylinder is incompressible. In such a case $\lambda \rightarrow \infty$ ($c=0$) and δ approaches zero, so that $\lambda\delta$ has a finite limit. (4) now takes the simple form

$$\beta^2 + (r\beta' + \beta)^2 = 2, \quad (5)$$

and

$$\lambda\delta = \mu[k + \int r\beta'^2 dr - \int \rho R dr], \quad (6)$$

where k is a constant.

From (5) we have

$$r\beta' = -\beta \pm \sqrt{2 - \beta^2}. \quad (7)$$

To decide the sign we observe that, since the radial displacement u_r is given by

$$u_r = (1 - \beta), \quad (8)$$

r_0 , the value of r before strain, is

$$r_0 = r\beta. \quad (9)$$

Hence

$$\frac{dr_0}{dr} = (r\beta' + \beta),$$

which must be positive unless the cylinder is turned inside out*. We should therefore take the positive sign with the under-root in (7).

Again

$$\frac{du}{dr} = 1 - (r\beta' + \beta) = 1 - \sqrt{2 - \beta^2},$$

which shows that, if u_r is positive and is to decrease with r , $\beta^2 < 1$. If u_r is negative, then $\beta^2 > 1$. $\beta=1$ will be seen to correspond to $r=\infty$. Also (9) shows that β should be positive.

Putting the value of $r\beta'$ from (7) in (6) we get

$$\lambda\delta = \mu \left[k + \frac{1}{2}\beta\sqrt{2 - \beta^2} - \frac{1}{2}\beta^2 + \sin^{-1} \frac{\beta}{\sqrt{2}} - \int \rho R dr \right]. \quad (10)$$

* *Loc. cit.* Proc. Roy. Soc. pp. 186-190.

If $0 < \beta < 1$ the solution of (7) is

$$\log \frac{C^2 \sqrt{2}}{r^2} = \log [\sqrt{2 - \beta^2} - \beta] + \tan^{-1} \frac{\sqrt{2 - \beta^2}}{\beta} \log_{10} c, \quad \dots (11.1)$$

all logarithms being to the base 10. C is a constant of integration.

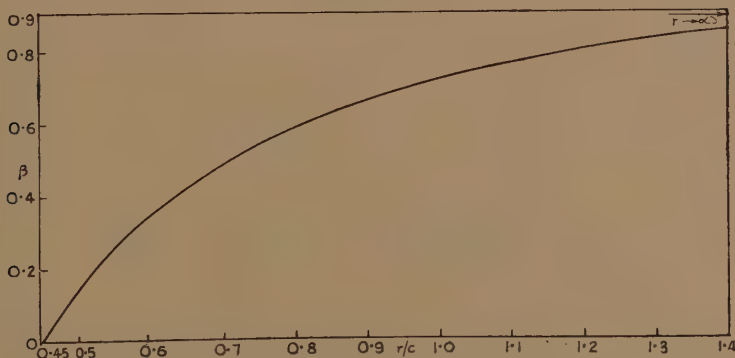
If $\sqrt{2} > \beta > 1$ the solution can be written as

$$\log \frac{C^2 \sqrt{2}}{r^2} = \log [\beta - \sqrt{2 - \beta^2}] + \tan^{-1} \frac{\sqrt{2 - \beta^2}}{\beta} \log_{10} c. \quad \dots (11.2)$$

TABLE I.

β	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
r/C	0.491	0.531	0.579	0.640	0.716	0.818	0.966	1.219	1.752	∞

Fig. 1.



We should use (11.1) or (11.2) according as the force is repulsive or attractive. In Table I. we have given the corresponding values of β and r/C as obtained from (11.1). In fig. 1 they have been plotted against each other.

To be able to proceed further we should know R . For a rotating shaft $\mu R = r\omega^2$, ω being the angular velocity. If the incompressible shaft is solid, $u_r = 0$, and the state of stress is that of a hydrostatic pressure. This needs no investigation. For a hollow shaft (10) gives

$$\lambda \delta = \mu \left[k + \frac{1}{2} \beta \sqrt{2 - \beta^2} - \frac{1}{2} \beta^2 + \sin^{-1} \frac{\beta}{\sqrt{2}} - \frac{1}{2} \frac{\rho}{\mu} r^2 \omega^2 \right]. \quad \dots (12)$$

Let the inner and outer boundaries of the cylinder be given by

$$\beta = 0.2 \quad \text{and} \quad \beta = 0.8.$$

The value of \widehat{rr} is

$$\widehat{rr} = \mu \left[k + \frac{1}{2} \beta^2 - 1 + \frac{1}{2} \beta \sqrt{2 - \beta^2} + \sin^{-1} \frac{\beta}{\sqrt{2}} - \frac{1}{2} \frac{\rho}{\mu} r^2 \omega^2 \right].$$

This is to vanish both over

$$r_1/C = 0.531 \quad (\beta = 0.2), \quad \text{and} \quad r_2/C = 1.210 \quad (\beta = 0.8).$$

Putting

$$C_1^2 = \frac{1}{2} \frac{\rho}{\mu} C^2 \omega^2,$$

we get the two equations

$$k + 0.381 = (1.464) C_1^2,$$

$$k - 0.698 = (0.282) C_1^2,$$

which give

$$k = 0.955, \quad C_1^2 = 0.913.$$

The non-vanishing stresses can now be written as

$$\widehat{rr} = \mu \left[0.955 + \frac{1}{2} \beta^2 - 1 + \frac{1}{2} \beta \sqrt{2 - \beta^2} + \sin^{-1} \frac{\beta}{\sqrt{2}} - (0.913) \left(\frac{r}{C} \right)^2 \right]. \quad (13.1)$$

$$\widehat{\theta\theta} = \mu \left[0.955 - \frac{3}{2} \beta^2 + 1 + \frac{1}{2} \beta \sqrt{2 - \beta^2} + \sin^{-1} \frac{\beta}{\sqrt{2}} - (0.913) \left(\frac{r}{C} \right)^2 \right]. \quad (13.2)$$

$$\widehat{zz} = \mu \left[0.955 - \frac{1}{2} \beta^2 + \frac{1}{2} \beta \sqrt{2 - \beta^2} + \sin^{-1} \frac{\beta}{\sqrt{2}} - (0.913) \left(\frac{r}{C} \right)^2 \right]. \quad (13.3)$$

The ordinary theory gives

$$u_{r_0} = \frac{1}{4} \frac{\rho \omega^2 r_1^2 r_2^2}{\mu r} = \frac{1}{2} C C_1^2 \frac{(r_1 r_2 / C)^2}{r / C}, \quad . \quad . \quad . \quad . \quad . \quad (14.1)$$

$$\widehat{rr}_0 = \mu C_1^2 \left[\left(\frac{r_1}{C} \right)^2 + \left(\frac{r_2}{C} \right)^2 - \left(\frac{r_1 r_2 / C}{r / C} \right)^2 - \left(\frac{r}{C} \right)^2 \right], \quad . \quad . \quad . \quad (14.2)$$

$$\widehat{\theta\theta}_0 = \mu C_1^2 \left[\left(\frac{r_1}{C} \right)^2 + \left(\frac{r_2}{C} \right)^2 + \left(\frac{r_1 r_2 / C}{r / C} \right)^2 - \left(\frac{r}{C} \right)^2 \right], \quad . \quad . \quad (14.3)$$

$$\widehat{zz}_0 = \mu C_1^2 \left[\left(\frac{r_1}{C} \right)^2 + \left(\frac{r_2}{C} \right)^2 - \left(\frac{r}{C} \right)^2 \right]. \quad . \quad . \quad . \quad . \quad . \quad (14.4)$$

Table II. gives the corresponding values of u_r , u_{r_0} , \widehat{rr} , \widehat{rr}_0 , etc.

TABLE II.

r/C .	\widehat{rr}/μ .	\widehat{rr}_0/μ .	$\widehat{\theta\theta}/\mu$.	$\widehat{\theta\theta}_0/\mu$.	\widehat{zz}/μ .	\widehat{zz}_0/μ .	u_r/C .	u_{r_0}/C .
0.531	0	0	1.92	2.67	0.96	1.34	0.43	0.68
0.579	0.12	0.17	1.94	2.41	1.03	1.29	0.41	0.56
0.640	0.22	0.30	1.90	2.14	1.06	1.22	0.38	0.46
0.716	0.30	0.39	1.80	1.86	1.05	1.13	0.36	0.37
0.818	0.35	0.42	1.63	1.54	0.99	0.98	0.33	0.28
0.964	0.30	0.34	1.02	1.15	0.81	0.74	0.29	0.20
1.210	0	0	0.72	0.51	0.36	0.26	0.24	0.13

Table II. shows that when the displacement is large the ordinary theory cannot be expected to give satisfactory results. As the displacement becomes smaller and smaller, the agreement between the two sets of values becomes closer and closer.

(b) *Spherical Shell.*

In this case the strain is again symmetrical, and we have

$$u=x(1-\beta), v=y(1-\beta), w=z(1-\beta),$$

where now β is a function of $r=(x^2+y^2+z^2)^{\frac{1}{2}}$ only.

The stresses are given by equations of the form

$$\widehat{xx}=\lambda\left[\frac{3}{2}(1-\beta^2)-\frac{1}{2}(r^2\beta'^2+2r\beta\beta')\right]+\mu\left[1-\beta^2-x^2\left(\beta'^2+\frac{2\beta\beta'}{r}\right)\right]. \quad (15.1)$$

$$\widehat{yz}=-\mu yz\left(\beta'^2+\frac{2\beta\beta'}{r}\right). \quad (15.2)$$

In order that they may satisfy the body-stress equations we find that β should satisfy the differential equation

$$2\beta^2+(r\beta'+\beta)^2+2c\int r\beta'^2 dr-c\int \rho R dr=K, \quad (16)$$

where K is a constant of integration, μR the radial body-force, and c has the same meanings as before.

Assuming that the shell is incompressible we get

$$2\beta^2+(r\beta'+\beta)^2=3, \quad (17)$$

and

$$\begin{aligned} \lambda\delta &= \lambda\left[\frac{3}{2}(1-\beta^2)-\frac{1}{2}(r^2\beta'^2+2r\beta\beta')\right] \\ &= \mu\left[k+2\int r\beta'^2 dr-\int \rho R dr\right]. \end{aligned} \quad (18)$$

From (17) we have

$$r\beta'=-\beta\pm\sqrt{3-2\beta^2}.$$

As before, the negative sign corresponds to the case of the shell turned inside out. Taking the positive sign we get

$$\lambda\delta=\mu\left[k-\beta^2+\beta\sqrt{3-2\beta^2}+3\sqrt{2}\sin^{-1}\frac{\beta\sqrt{6}}{3}-\int \rho R dr\right]. \quad (19)$$

When the force is repulsive β lies between 0 and 1, and the corresponding solution of (17) is

$$\log \frac{C^3\sqrt{3}}{r^3}=\log [\sqrt{3-2\beta^2}-\beta]+\left\{\sqrt{2}\tan^{-1}\frac{\sqrt{\frac{3}{2}-\beta^2}}{\beta}\log_{10} c\right\}, \quad (20.1)$$

all logarithms being to the base 10.

When it is attractive β lies between 1 and $\sqrt{\frac{3}{2}}$, and the solution of (17) becomes

$$\log \frac{C^3\sqrt{3}}{r^3}=\log [\sqrt{3-2\beta^2}-\beta]+\left\{\sqrt{2}\tan^{-1}\frac{\sqrt{\frac{3}{2}-\beta^2}}{\beta}\log_{10} c\right\}.$$

Tables III. and IV. give the corresponding values of β and r/C . In fig. 2 the values given in Table III. have been plotted.

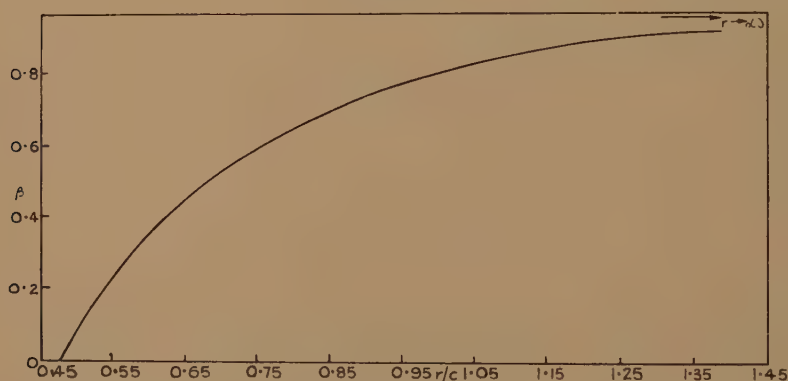
TABLE III.

β ..	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
r/C .	0.506	0.540	0.578	0.324	0.681	0.753	0.851	1.003	1.300	∞

TABLE IV.

β	1.225	1.20	1.16	1.12	1.10	1.08	1.04	1.0
r/C ...	1.049	1.151	1.218	1.324	1.391	1.487	1.833	∞

Fig. 2.



If we take the case of a self-gravitating shell, we have

$$\mu R = -\frac{4}{3}\pi\left(r - \frac{r_1^3}{r^2}\right),$$

r_1, r_2 being its inner and outer radii.

$\lambda\delta$ now becomes

$$\lambda\delta = \mu \left[k - \beta^2 + \beta\sqrt{3-2\beta^2} + 3\sqrt{2} \sin^{-1} \frac{\beta\sqrt{6}}{3} + \frac{4}{3}\pi \frac{\rho}{\mu} \left(\frac{1}{2}r^2 + \frac{r_1^3}{r} \right) \right]. \quad (20)$$

As the force is attractive we should use the values given in Table III. Taking $r_1/C = 1.151$ ($\beta = 1.2$) and

$$r_2/C = 1.833 (\beta = 1.04),$$

and putting

$$C_1^2 = \frac{4}{3}\pi \frac{\rho}{\mu} C^2,$$

the condition that $\widehat{r\dot{r}}$ is to vanish both at $r=r_1$ and $r=r_2$ gives

$$k + (1.988)C_1^2 + 2.760 = 0,$$

$$k + (2.511)C_1^2 + 2.184 = 0,$$

from which we have

$$C_1^2 = 1.102, \quad k = -4.951.$$

The non-vanishing stresses in spherical polar coordinates now take the form

$$\begin{aligned} \widehat{r\dot{r}} = \mu \left[-4.951 + \beta^2 - 2 + \beta\sqrt{3-2\beta^2} + 3\sqrt{2} \sin^{-1} \frac{\beta\sqrt{6}}{3} \right. \\ \left. + (1.102) \left\{ \frac{1}{2} \left(\frac{r}{C} \right)^2 + \frac{(r_1/C)^3}{r/C} \right\} \right], \quad \dots \quad (21.1) \end{aligned}$$

$$\begin{aligned} \widehat{\theta\dot{\theta}} = \phi\dot{\phi} = \mu \left[-4.951 - 2\beta^2 + 1 + \beta\sqrt{3-2\beta^2} + 3\sqrt{2} \sin^{-1} \frac{\beta\sqrt{6}}{3} \right. \\ \left. + (1.102) \left\{ \frac{1}{2} \left(\frac{r}{C} \right)^2 + \frac{(r_1/C)^3}{r/C} \right\} \right]. \quad \dots \quad (21.2) \end{aligned}$$

The ordinary theory gives

$$u_{r_0} = -(0.192)r \left(\frac{r_1}{r} \right)^3,$$

$$\widehat{r\dot{r}}_0 = \mu \left[-2.959 + (1.102) \left\{ \frac{1}{2} \left(\frac{r}{C} \right)^2 + \frac{(r_1/C)^3}{r/C} \right\} + (0.769) \left(\frac{r_1/C}{r/C} \right)^3 \right],$$

$$\widehat{\theta\dot{\theta}}_0 = \phi\dot{\phi}_0 = \mu \left[-2.959 + (1.102) \left\{ \frac{1}{2} \left(\frac{r}{C} \right)^2 + \frac{(r_1/C)^3}{r/C} \right\} - (0.384) \left(\frac{r_1/C}{r/C} \right)^3 \right].$$

Table V. gives the corresponding values of u_r , u_{r_0} , $\widehat{r\dot{r}}$, $\widehat{r\dot{r}}_0$, etc.

TABLE V.

r/C .	$-\widehat{r\dot{r}}/\mu$.	$-\widehat{r\dot{r}}_0/\mu$.	$-\widehat{\theta\dot{\theta}}/\mu$.	$-\widehat{\theta\dot{\theta}}_0/\mu$.	$-u_r/C$.	$-u_{r_0}/C$.
1.151	0	0	1.32	1.15	0.23	0.22
1.218	0.13	0.12	1.16	1.09	0.19	0.19
1.324	0.23	0.21	0.99	0.98	0.16	0.17
1.391	0.25	0.25	0.88	0.90	0.14	0.15
1.487	0.26	0.26	0.76	0.79	0.12	0.12
1.833	0	0	0.24	0.29	0.07	0.09

As the strain in this case is not very great, the agreement between the two sets of values is quite good.

Hindu College,
Delhi, India.

XXIX. *Factors affecting the Limit of Solubility of Elements
in Copper and Silver.*

By Professor E. A. OWEN, M.A., Sc.D., and
the late EDGAR WYNNE ROBERTS, M.Sc.,
University College of North Wales, Bangor *.

[Received October 29, 1938.]

NOTE.—The work described in this paper was in progress when the lamented death of Mr. Edgar Wynne Roberts took place on the 31st January, 1937. An interim report on the work was submitted to the Department of Scientific and Industrial Research on the 12th October, 1936.

1. INTRODUCTION.

ATTEMPTS have been made from time to time to correlate the results of the many investigations which have been carried out on binary alloy systems. One of the earliest advances in this direction was made by Hume-Rothery †, who discovered the electron concentration rule governing the appearance of intermediate phases in alloys of copper, silver, and gold. This rule was later found by Westgren ‡ and others to apply to many other alloy systems. Later Hume-Rothery § and his collaborators discussed the factors affecting the formation of α -solid solutions in silver and copper by the elements of the B sub-groups. From a survey of their experimental results they put forward certain conclusions which may be summarized as follows: (1) where the atomic diameters of solvent and solute atoms do not differ by more than a critical amount, the "size-factor" is favourable and considerable solid solution may take place, but if the size-factor is unfavourable, the range of solid solution is very restricted. (2) Where the size-factor is favourable, (a) the atomic depression of freezing-point is approximately proportional to the valency of the solute, or to a simple multiple or fraction of the normal valency; (b) when proceeding along the same horizontal row of the periodic table, the atomic depression of the melting-point is approximately proportional to the square of the valency; (c) the maximum α -solid solubility corresponds to a first approximation to a constant electron concentration.

* Communicated by Professor E. A. Owen.

† J. Inst. Metals, xxxv. p. 309 (1926).

‡ Metallwirtschaft, vii. p. 700 (1928).

§ Phil. Trans. Roy. Soc. A, cexxxiii. p. 1 (1934).

While these conclusions may be in general agreement with the facts, doubt is expressed by some writers * whether the whole number relations are exactly obeyed. Stockdale †, on the other hand, from an examination of the freezing-point depression produced by 5 atomic per cent. of the solute element found that, in nearly all cases, the whole number rule applies within the experimental error. Later work by Hume-Rothery ‡ showed that in the systems silver-cadmium and silver-tin, the relationship holds very closely.

Stockdale §, from a wide survey of experimental data, records other whole number relationships in binary alloys of copper and silver. He finds, for example, that in a saturated solid solution at the temperature of the peritectic horizontal, there is a simple integral relationship between the numbers of solvent and solute atoms. Recent work || on the silver-zinc system of alloys by X-ray analysis appears to support Stockdale's findings.

Hume-Rothery, Lewin, and Reynolds ¶ carried out a comprehensive series of accurate measurements on the lattice spacings of primary solid solutions of cadmium, indium, tin, and antimony in silver and of zinc, gallium, and germanium in copper. In dilute solid solutions in silver equal atomic percentages of cadmium, indium, tin, and antimony were found to expand the silver lattice by amounts proportional to 2 : 3 : 4 : 6 respectively. In dilute alloys of the copper series equal atomic percentages of zinc, gallium, and germanium expand the copper lattice by amounts proportional to the factors 3 : 4 : 4·8. The authors conclude that whilst whole number relations exist in both series, the effect of increasing valency is less marked in the copper series. In both series, however, increasing valency tends to expand the lattice, but the expansion is opposed by some factor which is relatively more important in the copper series.

The work described in this paper was in progress when these results were published. It was considered worth while to continue the work even if it only corroborated previous results, as the method of procedure was different from that employed by the above investigators. It had been planned to investigate in addition the effect of temperature. It is by no means obvious that if whole number relationships are found at one particular temperature they will still be the same at another temperature.

* Anon., 'Metallurgist,' October 1934.

† Proc. Roy. Soc. A, clii. p. 81 (1935).

‡ The Structure of Metals and Alloys. Inst. Metals Monograph, 1936.

§ *Loc. cit.*

|| Owen and Edmunds, J. Inst. Metals, lxiii. p. 297 (1938).

¶ Proc. Roy. Soc. A, clvii. p. 167 (1936).

In this paper an account is given of measurements made to investigate the effect of valency on the lattice parameters of silver and copper alloys in cases where the size-factor is favourable, and to find if temperature has any influence on the results. Use is made of the accurately determined parameter-composition curves to fix the α -phase boundaries of the silver alloys. Maximum α -solid solubility is also investigated in relation to lattice distortion and electron concentration.

2. EXPERIMENTAL.

Two focusing X-ray cameras, especially designed for high-temperature work, were employed during the investigation. These cameras differed only in minor details from the camera, which has already been described. The method of mounting, heating, and measuring the temperature of the powder specimens was also the same as previously*.

For silver alloys nickel radiation giving reflexions from the (422) planes was used. For copper alloys of parameter less than 3.63 Å. cobalt radiation giving reflexions from (400) planes for $K\alpha$ and from (331) planes for $K\beta$ radiation, but for alloys of parameter greater than 3.63 Å. the stronger reflexions from (331) planes with $K\alpha$ nickel radiation could be used.

The accuracy of the parameter measurements has previously been discussed†; values of parameter correct to ± 0.0003 Å. can be obtained from a single film for cubic lattices. All parameter values recorded in this paper have been corrected to 18° C., and they include the usual refractivity correction.

Wire-wound furnaces were used for annealing the alloys. For quenching experiments these furnaces could be operated from a storage battery of large capacity, and the temperatures could be accurately controlled. To measure the temperature an alumel-chromel thermocouple calibrated in the usual way with standard metals gave satisfactory results.

3. MATERIAL.

All the metals used were of a high degree of purity: copper, silver, cadmium, and zinc, 99.95; gallium and germanium, 99.9; indium and tin, 99.98; antimony and arsenic, 99.99 per cent. purity.

Silver-cadmium and copper-zinc alloys were prepared by heating weighed mixtures of fine filings of these metals in evacuated tubes of pyrex glass, at temperatures somewhat higher than the melting-points of cadmium and zinc respectively. A thorough investigation of this interdiffusion process had already been made with copper-zinc alloys‡.

* Owen and Yates, *Phil. Mag.* xvi. p. 606 (1933); xvii. p. 113 (1934).

† Owen and Yates, *Phil. Mag.* xv. p. 472 (1933).

‡ Owen and Pickup, *Proc. Roy. Soc.* cxlix. p. 282 (1935).

and the results showed that for complete diffusion the lattice parameters in the α -phase agreed with those obtained by the ordinary melting process. The method has also been used for alloys of zinc with silver *. In the present investigation samples of the silver-cadmium mixtures were removed at intervals for parameter measurement, and the diffusing temperature was progressively increased from 335° C. to 450° C., until it was definitely established that equilibrium had been attained. The copper-zinc mixtures were diffused at 500° C. and were also periodically examined.

This method of preparation gave alloys showing fine, intense reflexion lines, and since the material during the diffusion process was held in a confined space, there was no risk of loss of cadmium or zinc by volatilization, and the compositions of the alloys could be accurately determined from the weights of the constituent elements.

Apart from a few silver-tin alloys, which were prepared as ingots of about 20 gm. in salamander crucibles, all the remaining alloys were made in quantities not exceeding 3 gm. in silica tubes. The metals, in the form of small granules, were passed into silica tubes having a constriction near one end. After weighing, the tube was evacuated and sealed off at the constriction. The sealed tube was then attached to the arm of a shaking machine so that the metals, when molten in the furnace, could be thoroughly mixed. The rate of solidification of the alloy was either slow (by slowly cooling the furnace) or rapid (by quenching the shaking machine arm with tube attached). Ingots prepared in this way were in the form of small cylinders which were completely free from blow-holes and surface oxidation.

The ingots showed a loss in weight not exceeding 1 in 1000. The fact that silver-tin alloys showed the same loss suggested that the effect was not due to the volatilization of the solute metal. Pure silver showed about the same loss. From a spectroscopic analysis † of a silver ingot, prepared as described above, it was estimated that there was less than 0.008 per cent. of silicon present in the ingot. As only copper-rich and silver-rich alloys were prepared, most of the loss is in the copper or the silver, so there will be no appreciable error in composition from that deduced from the weights of the constituents. A glance at the figures in Table I., in which is recorded the compositions of a few alloys as deduced from the weights of the component elements and from chemical analyses of the alloy ingots ‡, will show that this is true.

* Owen and Pickup, *Proc. Roy. Soc. A*, *exl.* p. 344 (1933).

† We are indebted to Mr. D. M. Smith of the British Non-Ferrous Research Association for conducting this analysis.

‡ The chemical analyses were conducted by Mr. R. Johnston of the Midland Laboratory Guild (1928) Ltd., Birmingham.

As a further check a few silver-cadmium and copper-zinc alloys were prepared by the method just described, and their parameter values measured. These were found to fall exactly on the curves obtained with the alloys prepared by interdiffusion. It can therefore be assumed that the compositions of alloys prepared by this method can be obtained accurately from the weights of the constituent elements.

All ingots were given a preliminary annealing of about seven days at temperatures between 500° C. and 700° C. Particulars of these annealings for the various alloys are given later. The ingots were lump annealed before they were removed from the silica tubes in which they were prepared. In all cases the cylindrical ingots were cut with a fine saw along a diametrical plane. The sawings and filings taken from the

TABLE I.

Alloy.	Composition.	
	From weights of constituents.	From chemical analysis.
Cu-Zn .	8.01 % Zn ; 91.99 % Cu	7.98 % Zn ; 91.98 % Cu
Ag-Cd	7.98 % Cd ; 92.02 % Ag	8.10 % Cd ; 91.84 % Ag
Ag-In	4.84 % In ; 95.16 % Ag	5.03 % In * ; 94.99 % Ag
Ag-Sn	5.97 % Sn ; 94.03 % Ag	5.95 % Sn ; 93.94 % Ag
Ag-Zn	36.00 % Zn ; 64.00 % Ag	36.05 % Zn ; 63.98 % Ag

surface exposed by cutting were assumed to represent the mean composition of the ingot. In order to test for segregation, samples were taken from different parts of the ingot. It will be convenient to represent these parts by different letters, as shown in fig. 1. A, general mid-plane ; B, upper tip ; C, lower tip ; D, plane parallel to A about half-way between A and the circumference. These letters will be used in the tables of results.

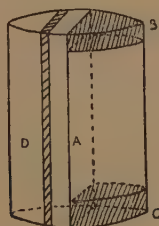
It will be shown later that in the silver alloys which were slowly cooled during solidification, considerable segregation occurred, and that quenching into cold water, on the other hand, produced too rapid cooling.

* Note by analyst : " We regard this result as perhaps a trace high because of the extremely hygroscopic nature of In_2O_3 . We have always found that, in spite of all the care we can give, this oxide tends to come high."

The procedure finally adopted was to quench all the silver alloys in oil from temperatures a few degrees above the liquidus temperatures. In addition the later alloys were made in quantities of 1 gm or 0.5 gm. In this way segregation effects were eliminated and considerable saving of material was effected.

All the copper alloys were prepared in 0.5 gm. ingots and were quenched in iced water to give rapid solidification.

Fig. 1.



I.—SILVER ALLOYS.

(A) *Determination of the Parameter-composition Curves.*

(a) *Pure Silver.*

Nine X-ray photographs of silver of 99.95 per cent. purity were taken with different radiations including those of nickel, cobalt, and copper. The mean value of the parameter at 18° C. was found to be 4.0772_2 \AA. , which is in excellent agreement with the value previously obtained with silver of this purity*.

Davey† found that the addition of minute traces of copper had a marked effect on the lattice parameter of silver. Silver of 99.999 per cent. purity had a parameter of $4.079 \pm 0.004 \text{ \AA.}$, but after adding copper to reduce the purity to 99.9 per cent., the parameter decreased to $4.058 \pm 0.004 \text{ \AA.}$ In view of this result Davey‡ comes to the conclusion that it is usually inadvisable to give lattice parameters to a greater accuracy than 0.1 per cent.

Since the chief impurity present in the silver used previously to determine the parameter of the material was copper, it was of interest to examine silver of a higher degree of purity. Material of 99.995 per cent. purity in crystal form was obtained which was rolled into a foil. This was annealed for one hour at 550° C. A foil of this kind probably yields parameter values of a slightly higher degree of accuracy than

* Owen and Yates, *Phil. Mag.* xv. p. 472 (1933).

† *Phys. Rev.* xxv. p. 753 (1925).

‡ ‘Study of Crystal Structure and its applications,’ 1934, p. 167.

those obtainable from a powder specimen. The mean of a number of determinations gave 4.0774_3 Å. as the parameter of silver of this degree of purity at 18°C . This value agrees well with the value 4.0773_7 given by Hume-Rothery, Lewin, and Reynolds *, and 4.0774_5 given by Jay †.

It appears therefore that the statement made by Davey is open to criticism and that the effects of small amounts of impurities are by no means so serious as he considers them to be.

(b) *Silver-Cadmium Alloys.*

Eleven alloys in the α -phase were prepared. Alloys with numbers 1 to 4, were 2 gm. mixtures, while the rest were each about 1 gm. The results obtained are given in Table II. It will be seen that the time of annealing required to produce equilibrium was rather longer for alloys of low cadmium content than for those richer in cadmium. Thus for alloy 8 containing 33.9 per cent. cadmium ‡, one week at 335°C . was sufficient,

TABLE II.
Lattice Parameters of Ag-Cd Alloys (α -phase)
prepared by Interdiffusion.

Heat treatment of powder mixture.	Alloy number and composition in at. % Cd.			
	1. 9.5_6	2. 19.1_9	3. 29.0_4	4. 38.9_8
2 weeks at 335°C	4.0973	4.1178	4.1415_5	4.1666
4 weeks at 335°C	4.0968	—	$\left\{ \begin{array}{l} 4.1412_5 \\ 4.1413_5 \end{array} \right\}$	—
$\left. \begin{array}{l} 2 \text{ weeks at } 335^\circ\text{C}. \\ + 2 \text{ weeks at } 380^\circ\text{C}. \end{array} \right\} \dots\dots$	—	4.1178	—	4.1666
$\left. \begin{array}{l} 4 \text{ weeks at } 335^\circ\text{C}. \\ + 2 \text{ weeks at } 380^\circ\text{C}. \end{array} \right\} \dots\dots$	4.0967	—	$\left\{ \begin{array}{l} 4.1415 \\ 4.1413_5 \end{array} \right\}$	—
$\left. \begin{array}{l} 4 \text{ weeks at } 335^\circ\text{C}. \\ + 2 \text{ weeks at } 380^\circ\text{C}. \\ + 1 \text{ week at } 450^\circ\text{C}. \end{array} \right\} \dots\dots$	4.0969	—	—	—
Mean values (Å.)	4.0968	4.1178	4.1414	4.1666

* Proc. Roy. Soc. A, clvii. p. 167 (1936).

† Z. Kristallog. lxxxvi. p. 106 (1933).

‡ Throughout this paper compositions of alloys are expressed in atomic percentages unless it is otherwise stated.

TABLE II. (cont.).

Heat treatment of powder mixture.	Alloy number and composition in at. % Cd.			
	5. 4.7 ₇	6. 14.2 ₈	7. 24.1 ₈	8. 33.9 ₀
1 week at 335° C.	4.0877	{ 4.1074 4.1075 }	4.1297 ₅	{ 4.1537 4.1535 }
3 weeks at 335° C.	4.0870 ₅	{ 4.1070 ₅ 4.1072 }	—	4.1535
1 week at 335° C. } +2 weeks at 380° C. }	—	—	4.1299 ₅	—
3 weeks at 335° C. +2 weeks at 380° C.	4.0867 ₅ 4.0870	} 4.1071 ₅	—	4.1535
3 weeks at 335° C. } +2 weeks at 380° C. }	4.0871	{ 4.1073 ₅ 4.1072 ₅ }	—	—
+1 week at 450° C. }				
Mean values (Å.)	4.0870	4.1072	4.1299	4.1535

Heat Treatment of powder mixture.	Alloy number and composition in at. % Cd.			
	9. 3.3 ₅	10. 7.7 ₈	11. 12.4 ₃	
4 weeks at 335° C. } +2 weeks at 380° C. }	4.0842	4.0932	{ 4.1026 ₅ 4.1026 ₅ }	
4 weeks at 335° C. } +2 weeks at 380° C. }	4.0841	4.0933 ₅	{ 4.1030 4.1030 }	
+1 week at 450° C. }				
Mean values (Å.)	4.0841 ₅	4.0933	4.1028	

but this was not so for alloy 5 containing 4.7₇ per cent. cadmium. This observation is in agreement with the result found * for the interdiffusion of copper and zinc that the phases in the surface layers of the copper particles are produced in order from the zinc-rich end, even when the composition of the mixture as a whole is in the pure α -region.

* Owen and Pickup, Proc. Roy. Soc. A, cxlix. p. 282 (1935).

Three silver-cadmium alloys were also prepared by fusion in silica. These were 0.5 gm. ingots, and were quenched in oil and lump annealed for seven days at 500° C. The parameter values obtained with these are given separately in Table III., and are found to agree closely with those obtained with the specimens prepared by interdiffusion.

TABLE III.

Lattice Parameters of Ag-Cd Alloys, prepared by Fusion.

Filings annealed 9 hrs. at 450° C.

No. of alloy.	Composition.	Lattice parameter Å.
1 A	5.0 ₀	4.0875 ₅
2 A	19.2 ₇	4.1184
3 A	34.7 ₉	4.1558

(c) *Silver-Indium Alloys.*

Eight alloys, seven in the pure α phase and one in the $(\alpha + \beta)$ region, were prepared. Of these numbers 1 to 4 were 3 gm. ingots and the remainder were 2 gm. ingots. All the ingots were allowed to cool very slowly over a period of several hours during solidification. The alloys were then sawn in the manner described without further treatment in lump form, and the sawings were given a heat treatment of 20 hours at 450° C. Later, when the texture of the lines indicated that the material was over-annealed, this time was reduced to 12 hours. Excellent reflexion lines were obtained.

The ingots were then sealed up in evacuated tubes and given a heat treatment of 7 days at 500° C. Samples were taken from different portions of some of the ingots as previously described. The results of the measurement of the films are given in Table IV.

It will be noticed that there is a certain amount of variation in composition in samples from different parts of the ingots. It was considered that sawings or filings from position A gave a good representative sample. There is, on the whole, good agreement between the values obtained from the sawings when no lump annealing was given and those obtained with filings from position A after seven days lump annealing at 500° C.

As a check on these values four more ingots were prepared. Alloy number 9 was a 1 gm. ingot and numbers 10, 11, and 12 were 0.5 gm.

ingots. These were all quenched in oil and lump annealed for seven days at 600° C. The results obtained with these alloys are given in Table V. When plotted these points and those obtained with alloys 1 to 7 fall on the same curve.

TABLE IV.
Lattice Parameters of Ag-In Alloys.

Alloy number and composition (at. % In).	Heat treatment of ingot.	Position of samples.				Estimated lattice parameter (Å.) at 18° C.
		A.	B.	C.	D.	
1. 2.44.	— 7 days 500°	4.0852 4.0851	— —	— —	— —	4.0851
2. 4.56.	— 7 days 500°	4.0916 ₅ 4.0916	— 4.0917	— 4.0913	— —	4.0916
3. 7.10.	{ — — 7 days 500°	4.0997	—	—	—	4.0997
		4.0996	—	—	—	
		4.0997 ₅	—	—	—	
4. 10.22.	— 7 days 500°	4.1102 4.1102	—	—	—	4.1102
5. 12.33.	{ — — 7 days 500°	4.1172	—	—	—	4.1173
		4.1174 ₅	—	—	—	
		4.1171 ₅	—	—	4.1174	
6. 15.14.	{ — — 7 days 500°	4.1270	—	—	—	4.1267
		4.1265	—	—	—	
		4.1267	4.1265	4.1262	4.1270	
7. 18.17.	— 7 days 500°	4.1378 4.1381	— —	— —	— —	4.1379

(d) *Silver-Tin Alloys.*

Four ingots, each of about 20 gm., were prepared by melting the metals under graphite in a salamander crucible lined with alundum cement. The melting was carried out in an atmosphere of nitrogen, and the molten alloy was well stirred with a graphite rod. The ingots were cast into sand moulds and were sawn into two pieces; one of these was lump annealed for three days at 600° C. and the other for 17 days at 600° C. Filings taken from different parts of the ingots yielded results which were too variable to be of any value, except with one alloy containing

3.58 per cent. tin which gave consistent parameter values, the mean of which was 4.0924.

The four alloys in the next batch were prepared in 3 gm. ingots in evacuated silica tubes and were slowly cooled during solidification. The filings were annealed for 20 hours at 450° C. The results are summarized in Table VI. A certain amount of variation in composition is to be seen also in these alloys, but the differences in parameter values are not great.

Three more ingots, each weighing about 3 gm., were prepared and quenched in iced water from a temperature of about 1000° C. The ingots were lump annealed for seven days at 700° C. and the filings given the

TABLE V.

Alloy number and composition (at. % In).	Filings from middle of ingot.	Filings from one end.	Estimated lattice parameter (Å.) at 18° C.
9. 16.23	4.1309 ₅ Å.	4.1307Å.	4.1309
10. 5.8	In these cases, the ingots were almost completely filed, and thus the parameter values were obtained with completely representative samples.		4.0960
11. 9.0			4.1064 ₅
12. 13.60			4.1218 ₅

same treatment as before. These and four other alloys prepared in 1 gm. ingots in narrow silica tubes and quenched in oil from temperatures just above the liquidus temperature, yielded more consistent results and are given in Table VII.

It will be noticed that the values of the parameters of these last alloys fall on the same curve as the parameters of the earlier ones. This shows that although there was variation of composition in the slowly cooled alloys, by taking filings from the general mid-plane (position A) a satisfactory representative sample was obtained.

(e) *Silver-Antimony Alloys.*

Six 3 gm. ingots were prepared using antimony of 99.8 per cent. purity. These were allowed to cool slowly during solidification. Sawings were taken without further lump annealing, and filings were

TABLE VI.

Lattice Parameters of Silver-Tin Alloys (α -phase).

Alloy number and composition (at. % Sn).	Heat treatment of ingot.	Position of samples.			Estimated lattice parameter (Å.) at 18° C.
		A.	B.	C.	
1 A. 2.73	— 3½ days 620° 8½ days 620°	4.0891 ₅	—	—	4.0892
		Unresolved lines giving parameter near that of pure silver (ingot oxidised).			
2nd half of ingot.	7 days 620°	—	4.0891 ₅	—	4.0893
	14 days 620°	—	—	4.0893	
2 A. 4.61	— 3½ days 620° 8½ days 620°	4.0969	—	—	4.0969
		—	4.0972	—	
		—	—	4.0966 ₅	
3 A. 6.36	— 1 day 620° 3½ days 620° 8½ days 620°	4.1046 ₅	—	—	4.1046
		—	—	4.1045	
		—	4.1049	—	
		4.1044 ₅	4.1048 ₅	—	
4 A. 9.17	— 3½ days 620° 8½ days 620°	4.1173 ₅	—	—	4.1174
		—	4.1173 ₅	—	
		—	—	4.1175 ₅	
2nd half of ingot.	7 days 620°	—	4.1175 ₅	—	4.1173
	14 days 620°	—	4.1174 ₅	4.1173	

TABLE VII.

Alloy no.	Composition (at. % Sn).	Lattice parameter (Å.) at 18° C.	Alloy no.	Composition (at. % Sn).	Lattice parameter (Å.) at 18° C.
6 A	1.71	4.0844	9 A	4.27	4.0954
7 A	5.57	4.1011	10 A	7.46	4.1098
8 A	7.06	4.1083	11 A	8.33	4.1133 ₅
			12 A	10.25	4.1219 ₅

taken after the ingots had been annealed for seven days at 550° C. The results are given in Table VIII.

TABLE VIII.

Alloy no.	Composition (at. % Sb).	Heat treatment.	Lattice parameter.	Mean lattice parameter (Å.) at 18° C.
1	1.79	— 7 days at 550°.	4.0880 4.0877 ₅	4.0879
2	3.59	— 7 days at 550° C.	4.0991 ₅ 4.0991	4.0991
3	5.41	— 7 days at 550° C.	4.1102 — *	4.1102
4	0.95	— 7 days at 550° C.	4.0830 4.0827 ₅	4.0829
5	2.66	— 7 days at 550° C.	4.0933 4.0931	4.0932
6	4.43	— 7 days at 550° C.	4.1044 4.1044 ₅	4.1044

TABLE IX.

Alloy no.	Composition (at. % Sb).	Mean lattice parameter (Å.) at 18° C.
1 A	1.14	4.0842
2 A	1.93	4.0888
3 A	2.70	4.0936
4 A	3.66	4.0995 ₅
5 A	4.80	4.1062
6 A	5.27	†4.1090 ₅
7 A	6.15	†4.1146 ₅

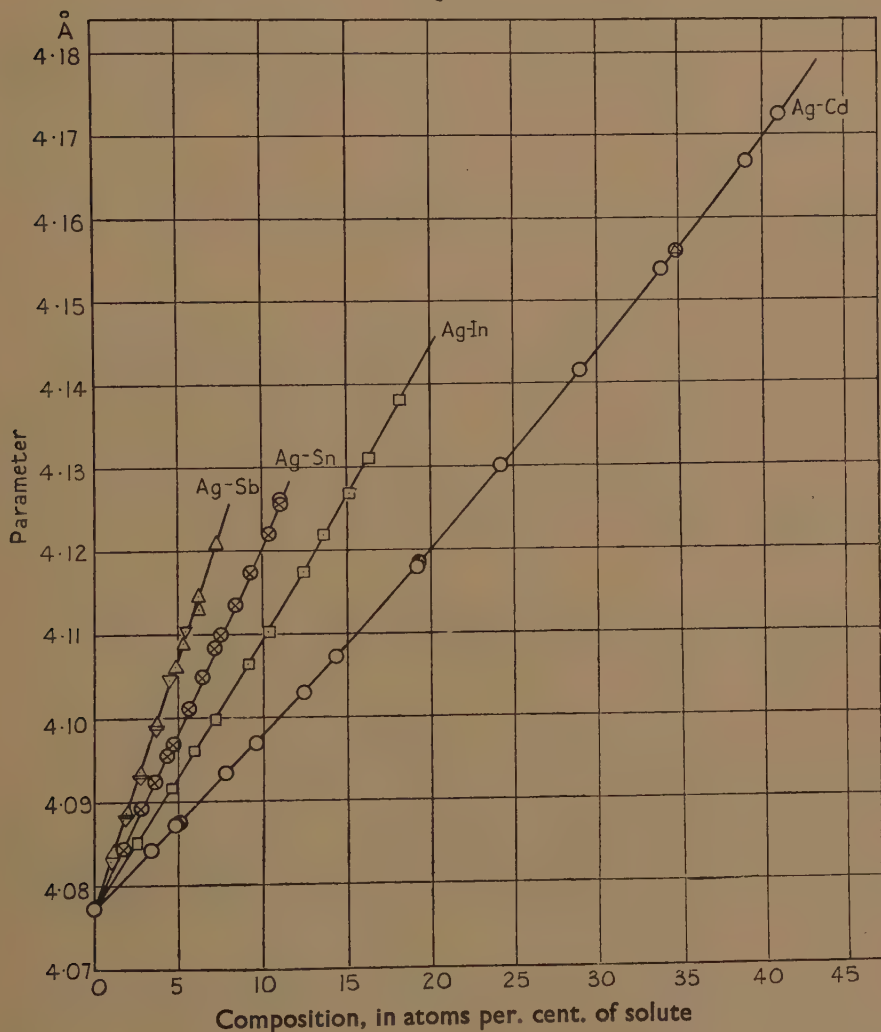
On plotting these values and comparing with the silver-cadmium curve, it was immediately evident that if a factor of 2 be assumed for cadmium, the factor for antimony was nearer to 6 than to 5. But before drawing any definite conclusions, it was decided to procure antimony of a higher degree of purity. "Kalbaum special" antimony, having a purity greater than 99.99 per cent., was obtained. With this material,

* The lines were too faint to measure with accuracy.

† Filings quenched after 2½ hours at 600° C.

seven alloys were prepared in 1 gm. ingots and quenched in oil from just above the freezing points. The ingots which were nearly spherical in form were lump annealed for 7 days at 600° C. The procedure adopted with these was to file one half of the ingot into powder and to take a

Fig. 2.



sample from this for measuring the parameter. The reflexion lines obtained were very fine. The results are shown in Table IX.

These values are shown by the points marked \triangle in fig. 2, and the curve was drawn with these points alone. It will be seen, however,

that the alloys made with antimony of 99.8 per cent. purity, represented by points marked ∇ , fall on the same curve. In this particular case the small amount of impurities present produces variations in the lattice parameters of silver-antimony alloys, which are within the limits of the accuracy of measurement.

All the parameter-composition curves are shown in fig. 2.

(B) *Determination of the α -phase Boundaries.*

As the parameter-composition curves for the pure α -phases had been carefully determined, it was considered worth while to determine the α -($\alpha+\beta$) boundaries, although in the case of cadmium, indium, and tin alloys these have been already investigated carefully by metallurgical methods by Hume-Rothery, Mabbott, and Channel-Evans.

(a) *Silver-cadmium boundary.*

A preliminary investigation was conducted to find whether there was serious loss of cadmium when samples were annealed in very small silica tubes at high temperatures. A sample of alloy 8 (33.9 per cent. cadmium) was annealed at 600° C. for 1 hour. The lattice parameter found after this heat treatment was 4.1536 Å., which agreed with the mean value 4.1535 Å. obtained after annealing at lower temperatures. As this alloy was in the pure α -phase, this agreement showed that there was no loss of cadmium. The fact that it was possible to prepare silver-cadmium alloys by fusion without loss of cadmium also shows that, provided the tubes are small, appreciable volatilization does not take place. It seems probable that Hume-Rothery and his collaborators, who found a loss of cadmium up to 1.5 per cent., used very much larger tubes. Those used in the quenching experiments described here were only about 10 mm. long with internal bore varying from zero up to about 2 mm. The amount of saturated vapour required to fill this cavity was negligibly small.

Four more alloys having compositions recorded in Table X. were prepared by diffusion for seven days at 400° C. The results of quenching experiments on these alloys and on alloy 4, together with the times of annealing, are included in the Table; the parameter values given are at 18° C.

In order to obtain efficient quenching a small vessel of iced water was pushed up into the furnace tube, the upper end being closed, to meet the dropped specimen tube; the specimen tube fell freely through about 2 or 3 cm. before entering the quenching medium. It was found that if this operation was carried out quickly there was no appreciable change in the temperature of the iron block in which the specimens were held during annealing, while the temperature of the quenching

medium remained constant. The silica tube containers had very thin walls. Below 500° C. it was found that thin-walled tubes of pyrex glass would stand quenching, and these were used instead of silica.

Above 600° C. it was found that the powder after annealing tended to sinter into a hard mass, and that, after breaking it up, the reflection lines were diffuse. This difficulty was overcome by mixing the powders with pure alumina powder for the annealing process. By using slightly different grades of alumina and alloy powders, they could be separated by sieves before taking the X-ray photograph.

TABLE X.

Details of annealing.	Alloy number and composition (at. % Cd).				
	4. 38·9 ₈ .	14. 41·0 ₄ .	12. 41·4 ₃ .	13. 43·0 ₃ .	15. 43·9 ₄ .
45 mins. 713° C.	4·1615	—	—	—	—
30 mins. 692° C.	4·1633 ₅	—	—	—	—
1 hr. 692° C.	4·1629	—	—	—	—
30 mins. 662° C.	—	4·1659	—	—	—
2 hrs. 600° C.	—	—	4·1691 ₅	4·1691 ₅	—
9½ hrs. 500° C.	—	—	—	4·1748	—
34 hrs. 500° C.	—	—	—	4·1746	4·1746 ₅
58 hrs. 500° C.	—	4·1724 ₅	—	4·1746	—
24 hrs. 485° C.	—	—	—	4·1752	—
39 hrs. 400° C.	—	—	—	4·1770 ₅	—
50 hrs. 400° C.	—	4·1723	—	—	4·1773 ₅
92 hrs. 300° C.	—	—	—	4·1763 ₅	—
110 hrs. 300° C.	—	4·1723	—	—	4·1766 ₅

The results given in Table X. were plotted and the boundary values obtained from the curves in the usual manner. These are shown in Table XI., with which are included for comparison the results of the only previous accurate investigation of this boundary which was carried out by Hume-Rothery, Mabbot, and Evans*.

(b) *Silver-Indium boundary.*

As this boundary is practically vertical, one alloy sufficed to determine it at all temperatures. The alloy chosen contained 20·9 atomic per cent. indium and was lump annealed for seven days at 500° C. Details of the annealing treatment given to the filings and the lattice parameter of

* Phil. Trans. Roy. Soc. A, cccxxiii. p. 1 (1934).

the specimens quenched from different temperatures are recorded in Table XII.

The compositions corresponding to the α -phase boundary obtained from these results are given, together with the values of Hume-Rothery, Mabbot, and Evans, in Table XIII. There is close agreement between the two determinations.

TABLE XI.
 α -Phase Boundary of Silver-Cadmium Alloys.

Temperature ° C.	Boundary (at. % Cd).	
	H-R, M and E.	O and R.
713	—	37.0
700	37.4	37.4
692	—	37.6
662	—	38.7
600	39.4	39.9
500	41.2	41.9
485	—	42.1
400	42.0	42.7 ₅
300	41.8	42.6

TABLE XII.

Details of annealing.	Lattice parameter at 18° C.
30 mins. 669° C.	4.1442 Å.
2 hrs. 643° C.	4.1434
2½ hrs. 600° C.	4.1425
9½ hrs. 500° C.	4.1421 ₅
24 hrs. 485° C.	4.1421 ₅
39 hrs. 400° C.	4.1417
92 hrs. 300° C.	4.1409 ₅

(c) *Silver-Tin Boundary.*

The results obtained by quenching samples of three silver-tin alloys which had been annealed in lump form for seven days at 600° C. are given in Table XIV., and the boundary derived from these results is recorded in Table XV. This table also contains the values of the compositions corresponding to the boundary obtained by combining the results of Hume-Rothery, Mabbot, and Evans with those of Murphy *.

* J. Inst. Metals, xxxv. p. 107 (1926).

(d) *Silver-Antimony Boundary.*

The results obtained with annealed specimens in the ($\alpha + \beta$) region of the silver-antimony system are given in Table XVI.

TABLE XIII.

α -Phase Boundary of Silver-indium Alloys.

Temperature ° C.	Boundary (at. % Indium).	
	H-R, M and E.	O and R.
669	—	19.9 ₅
650	19.8	19.7
643	—	19.6
600	19.6	19.5
500	19.4	19.4
400	19.4	19.3
300	19.4	19.1

TABLE XIV.

Details of annealing.	Alloy number and composition (at. % Sn).		
	12 A. 10.2 ₅ %.	5 A. 11.0 ₀ %.	14 A. 11.9 ₉ %.
20 mins. at 714° C.	—	—	4.1264 ₅
30 mins. at 700° C.	—	—	4.1266
1 hr. at 692° C.	4.1220	4.1259 ₅	—
30 mins. at 668° C.	—	—	4.1264 ₅
2½ hrs. at 600° C.	4.1219 ₅	4.1252	—
12 hrs. at 600° C.	—	4.1255	—
24 hrs. at 600° C.	—	—	4.1254 ₅
23 hrs. at 500° C.	4.1221 ₅	4.1239 ₅	—
54 hrs. at 500° C.	—	—	4.1241
9½ days at 400° C.	4.1219 ₅	4.1219 ₅	—
14 days at 300° C.	4.1207 ₅	—	—
19 days at 300° C.	4.1205	—	—

The values of the compositions at different temperatures along the boundary are given in Table XVII., together with the values obtained by Reynolds and Hume-Rothery*.

* J. Inst. Metals, lx. p. 365 (1937).

TABLE XV.

 α -Phase Boundary of Silver-Tin Alloys.

Temperature ° C.	Boundary (at. % Sn).	
	H-R, M and E, and Murphy.	O and R.
714	—	11.2
700	11.7	11.2
668	—	11.1 ₅
600	10.7	10.9
500	9.9 ₅	10.6
400	9.5	10.2
300	9.3	9.8

TABLE XVI.

Details of annealing.	Alloy number and composition (at. % Sb).		
	7 A. 6.1 ₅ %.	8 A. 7.2 ₀ %.	9 A. 7.9 ₇ %.
1 hr. at 700° C.	—	4.1209	4.1211 ₅
2½ hrs. at 600° C.	4.1146 ₅	—	—
12 hrs. at 600° C.	—	4.1190	4.1188 ₅
24 hrs. at 485° C.	—	4.1162	4.1159
68 hrs. at 400° C.	—	4.1143	4.1144
14 days at 300° C.	4.1113 ₅	4.1110	—

TABLE XVII.

 α -Phase Boundary of Silver-Antimony Alloys.

Temperature ° C.	Boundary (at. % Sb).	
	R and H-R.	O and R.
700	7.0	7.2
600	6.6	6.8 ₅
500	6.3	—
485	—	6.4
400	5.9 ₅	6.1
300	5.5 ₅	5.6

There is close agreement between the metallurgical and X-ray results for these alloys.

(C) *Thermal Expansion of the Silver Alloys.*

It was considered preferable to investigate the thermal expansion of a few alloys of each system by taking X-ray photographs at temperature intervals up to the maximum rather than to measure the parameters of all the alloys at one temperature only.

In order to check that the camera was giving correct results at high temperatures and that the thermocouple calibration was correct, X-ray photographs of silver were taken at intervals throughout the course of the investigation. In previous work on the thermal expansion of silver *, a nickel-nichrome thermocouple was employed. Although the results showed that this couple was capable of giving accurate temperature measurements, it is an inconvenient thermocouple to work with because its thermal e.m.f. curve shows certain characteristic irregularities between 250° C. and 400° C. An accurate calibration of the couple in this range is, therefore, difficult. The difficulty was overcome by comparing the nickel-nichrome couple with a platinum rhodio-platinum couple. For this purpose an electrically heated copper block of large mass was used with a long cylindrical hole bored along its axis, into which the two couples were placed with their junctions in contact. For most of the work described in this paper, however, an alumel-chromel thermocouple was used. This gave a straight line calibration curve over a wide range of temperature and was also found to stand exposure to high temperatures better than nickel-nichrome.

(a) *Thermal Expansion of Pure Silver.*

The parameter values of silver of purity 99.95 per cent. at temperatures up to 560° C. are given in Table XVIII. ; these values agree closely with the values found previously †.

(b) *Thermal Expansion of the Alloys.*

Exposures were made with several alloys of each system, the temperatures being chosen quite at random. The lattice parameter of each alloy at room-temperature was measured at frequent intervals to determine whether any loss of the solute element by volatilization had taken place. With cadmium and indium difficulty was experienced at temperatures above 300° C. because of this loss by volatilization.

* Owen and Yates, Phil. Mag. xvii. p. 113 (1934).

† Owen and Yates, Phil. Mag. xvii. p. 113 (1934).

Attempts were made to diminish the amount of volatilization by heating the specimen in an atmosphere of nitrogen instead of *in vacuo*. The nitrogen was purified by bubbling through tubes containing alkali solution of pyrogalllic acid, followed by calcium chloride, and trays of phosphorous pentoxide were placed in the chamber. With these precautions and using the nitrogen at about half an atmosphere pressure, exposures with silver-indium alloys at temperatures of about 330° C. were possible without appreciable change in the parameter. With

TABLE XVIII.

Thermocouple.	Temperature.	Lattice parameter (including refractivity correction).
Nickel-Nichrome	326° C.	4·1029 ₈ Å.
„	486	4·1170 ₈
„	398	4·1091 ₂
„	207	4·0926 ₉
Alumel-Chromel	401	4·1097 ₅
„	318	4·1021 ₆
„	313	4·1018 ₈
„	457	4·1145 ₅
„	557	4·1233 ₇
„	299	4·1005 ₇
„	135	4·0867 ₈
„	185	4·0910 ₅
„	97	4·0837 ₈

silver-tin and silver-antimony alloys no difficulty was experienced in making exposures at temperatures well above 400° C. In the case of alloy Ag-Cd 4, where the arcs from the (422) planes with nickel radiation were rather large at room-temperature, a vanadium target giving reflexions from (311) planes was used. In all other cases, nickel radiation was used.

The results collected in Tables XIX. to XXII. are entered in the table for each alloy system in the order in which the exposures were made.

(g) *Calculation of the Thermal Coefficients of Expansion.*

From the curves drawn by the aid of the figures in Tables XIX. to XXII. values of the lattice parameters at 300° C. for all the alloys were read off.

(c) *Silver-Cadmium Alloys*

TABLE XIX.

Composition (atomic per cent. Cadmium).							
7.7 ₈ .		14.2 ₈ .		29.0 ₄ .		39.9 ₈ .	
Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.
18	4.0933 ₅	18	4.1072 ₅	18	4.1413 ₅	18	4.1663
286	4.1159	169	4.1200 ₅	204	4.1576 ₅	259	4.1890
195	4.1080	286	4.1304 ₅	296	4.1667 ₅	18	4.1663 ₅
77	4.0982	217	4.1239	18	4.1414	156	4.1792
241	4.1119 ₅	145	4.1176	18	4.1399	196	4.1831
150	4.1041 ₅	97	4.1139 ₅	118*	4.1501 ₅	18	4.1665
115	4.1011 ₅	18	4.1075	75	4.1462	101	4.1738
18	4.0932	311	4.1324 ₅	258	4.1631	68	4.1707
325	4.1191	18	4.1075	163	4.1540 ₅	216	4.1851
371	4.1238	—	—	308	4.1674 ₅	293	4.1929
18	4.0927 ₅	—	—	—	—	18	4.1661
304*	4.1172 ₅	—	—	—	—	—	—
18	4.0934	—	—	—	—	—	—

(d) *Silver-Indium Alloys.*

TABLE XX.

Composition (atomic per cent. Indium).							
7.1 ₀ .		12.3 ₈ .		15.1 ₄ .		18.1 ₇ .	
Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.
88	4.1052	18	4.1171 ₅	18	4.1267	18	4.1379
164	4.1117 ₅	200	4.1321 ₅	92.5	4.1331 ₅	164	4.1503
18	4.0997 ₅	256	4.1372 ₅	139.5	4.1368 ₅	86	4.1435
135	4.1090 ₅	18	4.1172	18	4.1266	117.5	4.1462 ₅
214	4.1158	101	4.1241	†179.5	4.1402 ₅	204	4.1545
267.5	4.1202 ₅	144	4.1275	18	4.1267 ₅	18	4.1376
18	4.0995 ₅	226	4.1348 ₅	†284.5	4.1505 ₅	248	4.1589 ₅
299.5	4.1231 ₅	307	4.1424	18	4.1267 ₅	18	4.1378
18	4.0997	18	4.1168 ₅	†302	4.1528	289	4.1633
*18	4.0997	*18	4.1174 ₀	18	4.1263	18	4.1374 ₅
†326	4.1261	†322	4.1442	†266.5	4.1485	—	—
18	4.0996 ₅	18	4.1172 ₅	18	4.1262 ₅	—	—

* New specimen.

† Exposure made in nitrogen.

(e) *Silver-Tin Alloys.*

TABLE XXI.

Composition (atomic per cent. Tin).					
4·2 ₇ .		6·6 ₃ .		8·3 ₃ .	
Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.	Temp. ° C.	Lattice para- meter.
74	4·1000 ₅	18	4·1044 ₅	18	4·1133 ₅
121	4·1040 ₅	253	4·1240	300	4·1375
218	4·1120 ₀	313·5	4·1295 ₅	266	4·1346
180	4·1089 ₅	170	4·1165 ₅	220	4·1305 ₅
302	4·1192	125	4·1133	179	4·1265 ₅
18	4·0955 ₅	18	4·1044 ₅	122	4·1218
260	4·1155 ₅	209	4·1202	63	4·1172
333	4·1220 ₅	86	4·1097 ₅	18	4·1133
366	4·1252	290	4·1272	350	4·1420
396	4·1281 ₅	*18	4·1044 ₅	396	4·1464
18	4·0953	380	4·1359	18	4·1132
443	4·1327	334	4·1316 ₅	—	—
18	4·0955 ₅	412	4·1389	—	—
—	—	18	4·1044 ₀	—	—

(f) *Silver-Antimony Alloys.*

TABLE XXII.

Composition (atomic per cent. Antimony).			
2·7 ₀ .		4·8 ₀ .	
Temp. ° C.	Lattice parameter.	Temp. ° C.	Lattice parameter.
18	4·0935 ₅	18	4·1061 ₅
151	4·1045 ₄	293	4·1297 ₅
78	4·0988 ₅	18	4·1062 ₅
183	4·1073	193	4·1210 ₅
227	4·1110 ₅	236	4·1249 ₅
275	4·1152	266	4·1275
18	4·0938	18	4·1063
331	4·1199	331	4·1331 ₅
18	4·0937	18	4·1062
363	4·1237 ₅	136	4·1160
18	4·0937	86	4·1118 ₅
408	4·1278 ₅	364	4·1361 ₅
18	4·0937 ₅	18	4·1060 ₅
—	—	397	4·1392 ₅
—	—	18	4·1062
—	—	431	4·1429
—	—	18	4·1063

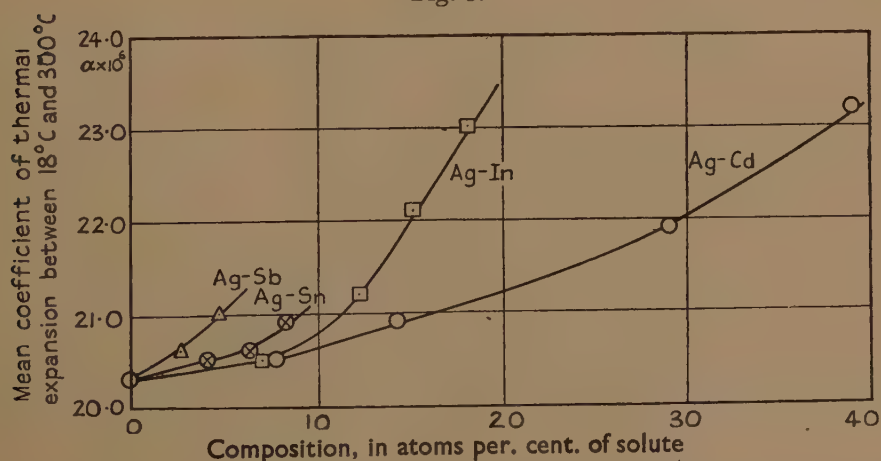
* New specimen.

These are included in Table XXIII. together with the values of the mean thermal coefficients between 18° and 300° C.

TABLE XXIII.

Element or alloy.	Atoms per cent. of solute.	$a_{18^{\circ}}$.	$a_{300^{\circ}}$.	$a_{300^{\circ}}-a_{18^{\circ}}$.	$\frac{a_{300^{\circ}}-a_{18^{\circ}}}{282 a_{18^{\circ}}}$.
Ag	—	40772 Å.	4·1006 Å.	·0234 Å.	$20\cdot3 \times 10^{-6}$
Ag-Cd	7·78	4·0933	4·1170	·0237	20·5
Ag-Cd	14·28	4·1074	4·1316	·0242	20·9
Ag-Cd	29·04	4·1414	4·1670	·0256	21·9
Ag-Cd	39·98	4·1664	4·1937	·0273	23·2
Ag-In	7·10	4·0997	4·1234	·0237	20·5
Ag-In	12·33	4·1172	4·1418	·0246	21·2
Ag-In	15·14	4·1267	4·1524	·0257	22·1
Ag-In	18·17	4·1378	4·1645	·0267	23·0
Ag-Sn	4·27	4·0955 ₅	4·1192	·0236 ₅	20·5
Ag-Sn	6·63	4·1044 ₅	4·1283	·0238 ₅	20·6
Ag-Sn	8·33	4·1133	4·1376	·0243	20·9
Ag-Sb	2·70	4·0937	4·1175	·0238	20·6
Ag-Sb	4·80	4·1062	4·1305	·0243	21·0

Fig. 3.



The change in the coefficient of expansion is not large, and the values of $a_{300^{\circ}}-a_{18^{\circ}}$ are subject to an error of at least ± 0.0002 Å. It is therefore

somewhat fortuitous that the values of the coefficients fall so well on smooth curves (fig. 3). After allowing for possible uncertainties in the values, however, the results show quite definitely that the change in the coefficient of expansion produced by a given atomic percentage of solute increases with increase of valency of the solute element.

The effect of the change in thermal expansion on the lattice parameter relationships will be discussed later.

II. COPPER ALLOYS.

(A) *Determination of Parameter-composition Curves.*

(a) *Pure Copper.*

The mean lattice parameter of copper of 99.9 per cent. purity was previously found to be 3.6077_5 Å at 18°C .^{*} Subsequent measurements of the parameter of this element made from time to time in this laboratory showed that the above value is somewhat high. For the purpose of this investigation material of higher purity (99.96 per cent.) was used, and the mean of several determinations of its parameter gave the value 3.6074_0 Å. at 18°C . This is still somewhat higher than the value recorded by Hume-Rothery, Lewin, and Reynolds †, namely, 3.6070_5 Å. at 18°C . We shall use the value found now (3.6074_0 Å.) rather than take the mean of the above determinations, because it is likely that the slight difference between the values is due to the different degrees of purity of the materials employed, and the alloys discussed in this paper were made with copper yielding this parameter value."

(b) *Copper-Zinc Alloys.*

In a previous investigation ‡ of the lattice parameters of the α -phase alloys of copper-zinc no alloys with less than 9 per cent. zinc were used. Also, although the method of analysis of the alloys was considered to give results correct to 0.15 per cent. zinc, some of the values showed rather large divergences from the smooth curve. Moreover, the parameter values were calculated from reflexion angles of 70° to 73° which are not in the most sensitive range of the focusing camera. It was therefore decided to establish the parameter-composition curve with greater accuracy, particularly in the region up to 10 atomic per cent. zinc.

Five alloys were prepared by interdiffusion as previously described, and three ingots were prepared by melting the metals in silica containers. The latter were given a preliminary annealing of eight days at 500°C .

^{*} Owen and Yates, *Phil. Mag.* xv. p. 472 (1933).

[†] *Loc. cit.*

[‡] Owen and Pickup, *Proc. Roy. Soc. A*, cxxxvii. p. 397 (1932).

They were half-gramme ingots and were completely reduced to filings which were therefore completely representative of the compositions of the alloys.

The results are given in Tables XXIV. and XXV.

TABLE XXIV.

Lattice Parameters at 18° C. of Cu-Zn Alloys,
prepared by Interdiffusion.

Heat treatment.	Alloy number and composition (at. % Zn).				
	1. 3·6 ₄ %.	2. 7·3 ₂ %.	3. 14·7 ₈ %.	4. 24·4 ₅ %.	5 34·2 ₅ %.
6 days at 500°	3·6145 Å.	3·6227 Å.	3·6390 Å.	—	—
12 days at 500°	3·6146 ₅	3·6225	3·6382	—	—
23 days at 500°	3·6144	3·6222 ₅	3·6381 ₅	—	—
15 days at 500°	—	—	—	3·6610 ₅	3·6849

TABLE XXV.

Lattice Parameters of Cu-Zn Alloys, prepared by Fusion.
Filings annealed 8 hrs. at 400° C.

Alloy number and composition (at. % Zn).	Lattice para- meter (Å.) at 18° C.
1 A. 5·9 ₆	3·6193 ₅
2 A. 11·8 ₀	3·6317 ₅
3 A. 19·8 ₀	3·6505

(c) Copper-Gallium Alloys.

These alloys were prepared in half-gramme ingots by melting, quenched in water for rapid solidification, and lump annealed for 10 days at 600° C. Filings were taken alternately from the upper and lower tips of consecutive ingots, so that if there was any variation in composition, the

Fig. 4.

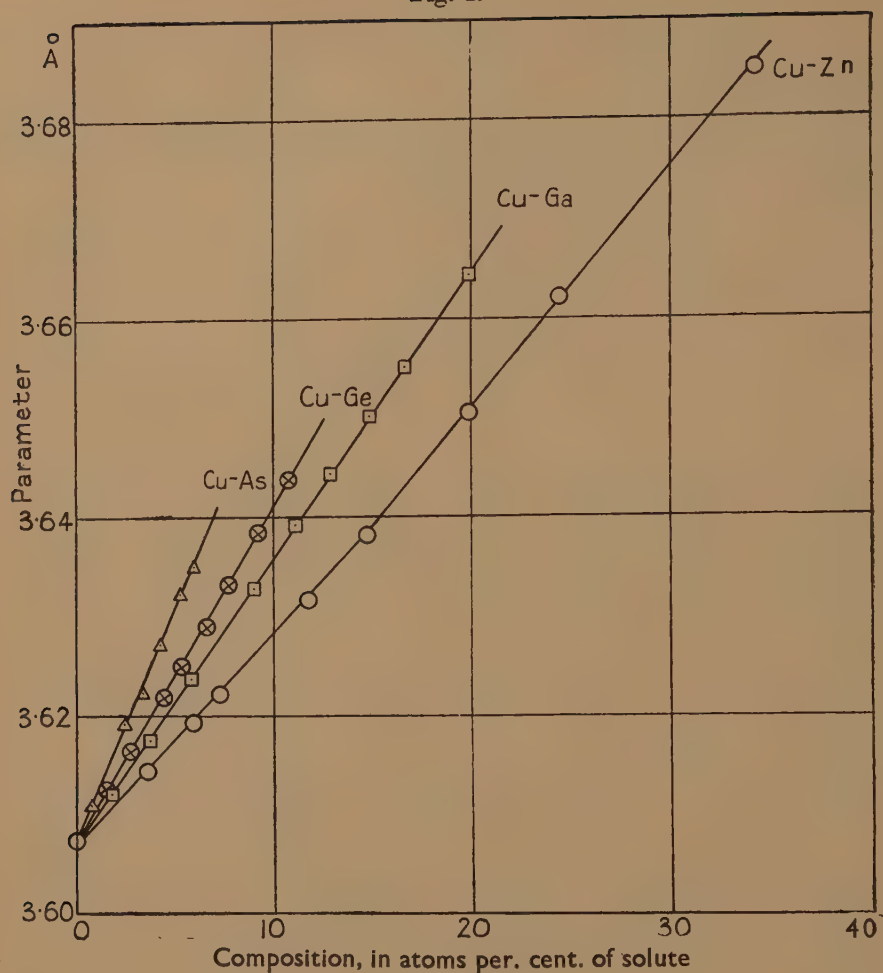


TABLE XXVI.

Filings annealed 10 hours at 450°.

Alloy number.	Composition (at % Ga).	Lattice parameter at 18° C.
1	11.1 ₇	3.6392 ₅ Å.
2	1.7 ₉	3.6121 ₅
3	3.7 ₀	3.6177
4	5.8 ₄	3.6238
6	8.9 ₇	3.6328 ₅
7	12.9 ₁	3.6444
8	14.8 ₇	3.6501 ₅
9	16.6 ₄	3.6551
10*	19.9 ₀	3.6643 ₅

* Filings annealed 3 hrs. at 600° and quenched.

points should be arranged progressively above and below the mean curve. It will be seen from fig. 4, however, that a smooth curve can be drawn through all the points.

(d) *Copper-Germanium Alloys.*

These alloys were treated in exactly the same manner as the copper-gallium alloys. The results are given in Table XXVII.

TABLE XXVII

Alloy number.	Composition (at. % Ge).	Lattice parameter at 18° C.
1	7.7 ₈	3.6332 Å.
2	1.5 ₃	3.6125
3	2.7 ₅	3.6164 ₅
4	4.4 ₁	3.6219
5	5.3 ₅	3.6250 ₅
6	6.6 ₇	3.6291
7	9.2 ₅	3.6384 ₅
8*	10.8 ₁	3.6438

(e) *Copper-Arsenic Alloys.*

The following table gives the lattice parameter and composition of the α -phase copper-arsenic alloys.

TABLE XXVIII.

Alloy number.	Composition (at. % As).	Lattice parameter at 18° C.
1	0.7 ₉	3.6110 ₅ Å.
3	2.4 ₇	3.6193 ₅
4	3.3 ₈	3.6225
5	4.2 ₂	3.6273 ₅
6	5.2 ₉	3.6324 ₅
7	5.9 ₄	3.6353

III. DISCUSSION OF RESULTS.

(a) *Effect of Valency of Solute on Distortion produced in the Lattice of the Solvent.*

In order to investigate the effect of the valency of the solute elements on the distortion produced in the silver and copper lattices, the values

* Filings annealed $\frac{1}{2}$ hour at 700° and quenched.

TABLE XXIX.

Concentration of solute element.	Alloy system.	Lattice parameter.	Increase from pure metal.	Factor at 18° C.
3 % ...	Ag-Cd	4·0833	·0061	2·0
	Ag-In	4·0868	·0096	3·14
	Ag-Sn	4·0902	·0130	4·25
	Ag-Sb	4·0955	·0183	5·99
6 % ...	Ag-Cd	4·0895	·0123	2·0
	Ag-In	4·0963	·0191	3·11
	Ag-Sn	4·1032	·0260	4·23
	Ag-Sb	4·1138	·0366	5·95

Lattice parameter.	Alloy system.	Composition (atoms per cent. of solute).	Factor.
4·0960..	Ag-Cd	9·15	2·0
	Ag-In	5·91	3·09
	Ag-Sn	4·35	4·21
	Ag-Sb	3·10	5·90

Concentration of solute element.	Alloy system.	Lattice parameter.	Increase from pure metal.	Factor at 18° C.
3 % ..	Cu-Zn	3·6133	·0059	3·0
	Cu-Ga	3·6157	·0083	4·22
	Cu-Ge	3·6173	·0099	5·03
	Cu-As	3·6214	·0140	7·11
6 % ..	Cu-Zn	3·6194	·0120	3·0
	Cu-Ga	3·6241	·0167	4·16
	Cu-Ge	3·6273	·0199	4·96
	Cu-As	3·6356	·0282	7·04

Lattice parameter.	Alloy system.	Composition (atoms per cent. of solute).	Factor.
3·6220 ..	Cu-Zn	7·30	3·0
	Cu-Ga	5·25	4·17
	Cu-Ge	4·43	4·95
	Cu-As	3·12	7·02

of the lattice parameters of the solid solutions for concentrations of 3 and 6 atomic per cent. of solute, and also the concentrations of the solute metals required to produce the same lattice distortion, have been taken from the curves in figs. 2 and 4. These are entered in Table XXIX., and from these are calculated the relative magnitude of the changes produced by the different solute elements. The parameters of the silver and copper used to make the alloys are taken to be 4.0772 Å. and 3.6074 Å. respectively.

The different sets of values with both the silver and copper alloys yield results which are in good agreement. The mean values are included in Table XXX.

In the silver series the lattice distortions produced per atom of cadmium, indium, tin, and antimony are approximately in the ratio 2 : 3 : 4 : 6, as already found by Hume-Rothery, Lewin, and Reynolds.

TABLE XXX.

Silver alloys.		Copper alloys.	
Solute element.	Factor.	Solute element.	Factor.
Cd	2.0	Zn	3.0
In	3.11	Ga	4.18
Sn	4.23	Ge	4.98
Sb	5.95	As	7.06

The deviation from these figures is greatest in the case of the silver-tin alloys, the departure being about 6 per cent. This is just outside the limits of accuracy given by Hume-Rothery and his collaborators. Apart from this figure, the agreement between the present values and theirs is satisfactory.

The case of the copper alloys is more satisfactory in that the maximum deviation from the ratios 3 : 4 : 5 : 7 for the lattice distortion per atom of zinc, gallium, germanium, and arsenic, amounts only to about 4 per cent., which occurs for the copper-gallium alloys. We regard this deviation as within the amount that can be tolerated. These results with the copper alloys differ from those found by Hume-Rothery and his collaborators. They found that the lattice distortions per atom of zinc, gallium, and germanium vary as 3 : 4 : 4.8. Having regard to the possible sources of error in an investigation of this nature, it is reasonable to expect deviations amounting to 5 per cent. in the final figures.

It may be stated definitely that the valency of the solute element has an effect on the distortion produced, but the effect is not proportional to the valency.

Hume-Rothery's view is that when the sizes of the solute and solvent atoms are within certain critical limits and when the elements are near in the periodic table, only valency effects need be considered. The deviations in the results now obtained suggest that the size factor and/or some other factor may be operating even when the difference in the sizes of the atoms is relatively small.

(b) *Effect of Temperature.*

The thermal expansion values obtained with the silver alloys enable the factors to be evaluated at 300° C. This has been done for the parameter values of alloys containing 6 atomic per cent. of solute. The method of calculation will be clear from Table XXXI.

TABLE XXXI.

	Lattice parameter at 18° C.	Mean coef. of expansion. $\alpha_{(18-300)}$	$a_{300}-a_{18}$	a_{300}	Increase from pure metal.	Factor at 300° C.
Ag	4.0772	20.3×10^{-6}	0.0233 ₅	4.1005 ₅	—	—
Ag-Cd	4.0895	20.5	0.0236 ₀	4.1131 ₀	0.0125 ₅	2.0
Ag-In	4.0963	20.5	0.0236 ₅	4.1199 ₅	0.0194 ₀	3.09
Ag-Sn	4.1032	20.6	0.0238 ₅	4.1270 ₅	0.0265 ₀	4.21
Ag-Sb	4.1138	21.2	0.0246 ₀	4.1384 ₀	0.0378 ₅	6.03

The factors in the last column agree with those calculated from parameters measured at room-temperatures, so that the effect of thermal expansion on the factors is inappreciable. The effect of temperature required investigation because of the large differences existing between the coefficients of thermal expansion of the elements cadmium, indium, tin, and antimony; for example, the mean linear coefficient of expansion of cadmium * between 18° C. and 300° C. is 37×10^{-6} , whereas that of antimony † over the same range of temperature is 10.3×10^{-6} . If the coefficient of thermal expansion of an alloy were given approximately by a linear relation between the coefficients of its constituent metals, it was calculated that if the parameter factor for antimony were 6 at 18° C., it would be about 5.5 at 300° C. The results of the expansion measurements show, however, that the change in the thermal coefficient

* Owen and Roberts, Phil. Mag. xxii. p. 290 (1936).

† Hidnert, J. Research Nat. Bur. Standards, xiv. p. 523 (1935).

produced by the addition of the solute elements is very much less than that calculated on the above assumption, and that antimony which has a much smaller expansion than the other elements, causes the largest increase in the expansion of the silver lattice.

From the above measurements at high temperatures it is concluded that the parameter factors are approximately independent of temperature.

(c) *Maximum Solubility and Lattice Distortion.*

From the measurements made to determine the α -phase boundaries, the figures contained in Table XXXII. were obtained for the limits of solubility of the various elements in silver and copper respectively, and for the lattice parameters at these limits. The data concerning

TABLE XXXII.

Alloy system.	Maximum α -solid solubility (atoms/cent.).	Temp. at which maximum occurs.	Lattice parameter at max. solution. (a)	Atomic volume of solute. (v)	$\frac{100}{v} \cdot \frac{\delta a}{a}$
Ag-Cd	42.8 Cd	440° C.	4.1772 Å.	13.02	0.187 ₃
Ag-In	20.2 In	692	4.1447	15.72 ₅	0.105 ₁
Ag-Sn	11.2 Sn	724	4.1265	16.31	0.073 ₈
Ag-Sb	7.2 Sb	702.5	4.1219	18.24	0.059 ₉
Cu-Zn	38.4 Zn	450	3.6965	9.16 ₈	0.269 ₄
Cu-Ga	19.7 Ga	620	3.6641	11.76 ₇	0.133 ₄
Cu-Ge	11.4 Ge	821	3.6459	13.65 ₇	0.078 ₁

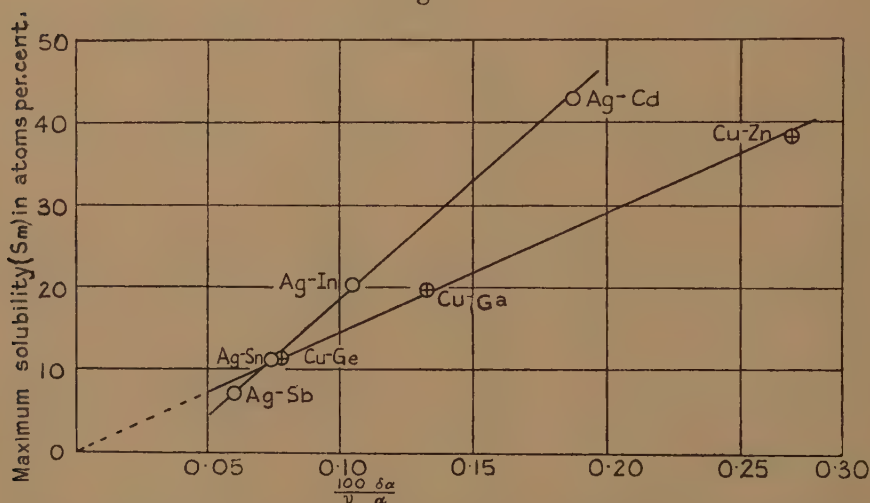
the copper alloys are based on the results of another investigation which has not yet been published. The temperatures at which maximum solubility occurs are those of the peritectic horizontals or the eutectic transformations, and are taken from the most recent investigations of the various alloy systems concerned. The table contains also the atomic volume of the solute atoms (*i. e.*, the atomic weight divided by the density), and the percentage lattice distortion divided by the atomic volume of the solute atom. Plotting this quantity against the maximum solubility in atoms per cent. of the solute element (fig. 5), an approximately linear relation is found to exist between the two quantities. There appears to be direct proportionality between the quantities in the case of alloys in the copper series, so that the maximum lattice distortion in the α -phases of alloys in this series is directly proportional to the total

volume of the atoms of solute entering into solution. With the silver alloys the maximum lattice distortion is directly proportional to the product of the atomic volume and the maximum solubility plus a constant. The results may be represented by the relation

$$\delta a/a = bv(s_m + c),$$

where $\delta a/a$ is the proportional lattice distortion, v the atomic volume of the solute as defined above, s_m the maximum solubility in atoms per cent. of solute, b and c are constants, each having different values for different alloy series. The present results show that for the copper series $c=0$, and for the silver series $c=10$. The copper results refer only to the elements zinc, gallium, and germanium. Alloys with the element arsenic have to be examined further.

Fig. 5.



It is interesting to note that whereas for copper-zinc, copper-gallium, and copper-germanium alloys the maximum value of the proportional distortion of the α -phase lattice bears a constant ratio to the volume of the solute producing the distortion, a result which might have been anticipated, for the silver-cadmium, silver-indium, silver-tin, and silver-antimony alloys, the maximum value of the proportional distortion bears a constant ratio to a quantity $(s_m + c)v$, which is greater than the volume of the solute present.

(d) *Maximum Solubility and Electron Concentration.*

The electron concentration corresponding to the maximum solubility has been calculated for the two series of alloys. The values are included in Table XXXIII.

In accordance with the limits set down by Hume-Rothery, the atomic size-factors in these alloys are favourable. In each series of alloys the electron concentration (*i. e.*, the ratio of valency electrons to atoms) corresponding to maximum α -solution, shows a definite decrease, which is more regular with the silver alloys than with the copper alloys. The total decrease amounts to about 10 per cent. in each case, but the figures for the copper-arsenic alloys are not so reliable as those for the other alloys.

The present results show in general that as the electron concentration diminishes, the percentage distortion also diminishes, but at a proportionately greater rate. The investigation needs to be extended to

TABLE XXXIII.

Alloy system.	Maximum α -solution (atoms per cent.).	Electron concentration at maximum.
Ag-Cd	42.8	1.42 ₈
Ag-In	20.2	1.40 ₄
Ag-Sn	11.2	1.33 ₈
Ag-Sb	7.2	1.30 ₈
Cu-Zn	38.3	1.38 ₈
Cu-Ga	19.7	1.39 ₄
Cu-Ge	11.4	1.34 ₂
Cu-As	6.8	1.27 ₂

cover more alloys before attempting to find a relation between the electron concentration and other quantities which it may influence.

SUMMARY.

The first part of the paper deals with the accurate determination of the crystal parameters of primary solutions of cadmium, indium, tin, and antimony in silver, and of zinc, gallium, germanium, and arsenic in copper. The boundaries of the α -phases have been determined for the silver alloys by the X-ray method.

The latter part of the paper deals with (*a*) the effect of the valency of the solute atom on the distortion produced in the lattice of the solvent, (*b*) the effect of temperature on the parameter factor, (*c*) the relation between maximum α solid solubility and lattice distortion, (*d*) the relation between maximum solubility and electron concentration.

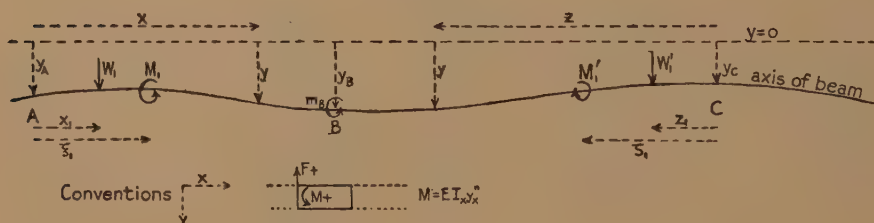
XXX. The Theorem of Three Moments with Variable "I" but without Thrust.

By F. J. TURTON, B.Sc., A.R.C.Sc.,
City and Guilds College *.

[Received October 24, 1938.]

§ 1. IN this paper the author's new form of the Theorem of Three Moments†, giving the relation between the bending moments at any three consecutive supports A, B, C for a continuous beam having *any* distributed load w per unit length at X from A in AB, any number of concentrated loads W_1 at x_1 ,

Fig. 1.



W_2 at x_2 etc. from A in AB, and any applied couples M_1 at ξ_1 , M_2 at ξ_2 etc. from A in AB, and similar loads and couples on CB, together with couples m_A at A, m_B at B etc., but without thrust or tension, is extended to *continuous* variation of the second moment of the section in the bays,

in the case where the second moment is $I_x = I_A \left(1 + \sum_{n=1}^{n=s} K_n x^n / a^n \right)$ at x from

A in AB and is $I_z = I_C \left(1 + \sum_{n=1}^{n=t} K_n' z^n / c^n \right)$ at z from C in CB, s and t being

any positive integers, a and c the lengths of AB and CB respectively, and the K 's and K' 's numerical constants.

* Communicated by Dr. W. G. Bickley.

† Phil. Mag. (7) xxv. p. 852 (1938). The notation of that paper is used in the present extension.

With the conventions of fig. 1 and omitting for the moment the distributed load and the applied couples in the bay, we have, if $x > x_1, x_2$ etc., but $\leq a$,

$$EI_x y'' = M_{AR} - x F_{AR} + \Sigma W_1 (x - x_1)$$

$$EI_A y'' = \left(1 + \sum_{n=1}^s K_n x^n / a^n\right) M_{AR} - \left(x + \sum_{n=1}^s K_n x^{n+1} / a^n\right) F_{AR} \\ + \Sigma W_1 \left((x - x_1) + \sum_{n=1}^s K_n (x^{n+1} - x^n x_1) / a^n \right).$$

Integrating by Macaulay's method*, terms in W_r are integrated from x_r to x , giving

$$EI_A y' = EI_A y'_{AR} + \left(x + \sum_{n=1}^s \frac{K_n x^{n+1}}{(n+1)a^n}\right) M_{AR} - \left(\frac{x^2}{2} + \sum_{n=1}^s \frac{K_n x^{n+2}}{(n+2)a^n}\right) F_{AR} \\ + \Sigma W_1 \left(\frac{(x - x_1)^2}{2} + \sum_{n=1}^s \frac{K_n}{a^n} \left(\frac{x^{n+2}}{n+2} - \frac{x^{n+1} x_1}{n+1} + \frac{x_1^{n+2}}{(n+1)(n+2)} \right) \right), \\ EI_A y = EI_A y_A + x EI_A y'_{AR} + \left(\frac{x^2}{2} + \sum_{n=1}^s \frac{K_n x^{n+2}}{(n+1)(n+2)a^n}\right) M_{AR} \\ - \left(\frac{x^3}{2 \cdot 3} + \sum_{n=1}^s \frac{K_n x^{n+3}}{(n+2)(n+3)a^n}\right) F_{AR} \\ + \Sigma W_1 \left(\frac{(x - x_1)^3}{2 \cdot 3} + \sum_{n=1}^s \frac{K_n}{a^n} \left(\frac{x^{n+3}}{(n+2)(n+3)} - \frac{x^{n+2} x_1 - x_1^{n+2} x}{(n+1)(n+2)} \right) \right).$$

At $x = a$, we get

$$EI_{BY_{BL}}'' = M_{BL} = M_{AR} - a F_{AR} + \Sigma W_1 (a - x_1), \quad \dots \quad (1)$$

$$EI_A y'_{BL} = EI_A y'_{AR} + a \left(1 + \sum_{n=1}^s \frac{K_n}{n+1}\right) M_{AR} - a^2 \left(\frac{1}{2} + \sum_{n=1}^s \frac{K_n}{n+2}\right) F_{AR} \\ + \Sigma W_1 \left(\frac{(a - x_1)^2}{2} + \sum_{n=1}^s \frac{K_n}{a^n} \left(\frac{a^{n+2}}{n+2} - \frac{a^{n+1} x_1}{n+1} + \frac{x_1^{n+2}}{(n+1)(n+2)} \right) \right), \quad (2)$$

$$EI_A y_B = EI_A y_A + a EI_A y'_{AR} + a^2 \left(\frac{1}{2} + \sum_{n=1}^s \frac{K_n}{(n+1)(n+2)}\right) M_{AR} \\ - a^3 \left(\frac{1}{2 \cdot 3} + \sum_{n=1}^s \frac{K_n}{(n+2)(n+3)}\right) F_{AR}$$

* For the same constant of integration in y' and y to apply throughout $0 \leq x \leq a$, the integrals arising from W_r (applied at x_r) must vanish at x_r , y' and y being continuous at x_r . This is achieved by integrating these terms from x_r to x , the method of W. H. Macaulay, 'Messenger of Mathematics,' xlviii. p. 119. When the term concerned is $A(x - x_r)^p$, A being constant, the rule gives

$$\int_{x_r}^x A(x - x_r)^p dx = A \int_0^{x - x_r} (x - x_r)^p d(x - x_r) = A(x - x_r)^{p+1} / (p+1),$$

a less general form to which the phrase "Macaulay's method" is frequently applied.

$$+ \Sigma W_1 \left(\frac{(a-x_1)^3}{2 \cdot 3} + \sum_{n=1}^s \frac{K_n}{a^n} \left(\frac{a^{n+3}-x_1^{n+3}}{(n+2)(n+3)} - \frac{a^{n+2}x_1-ax_1^{n+2}}{(n+1)(n+2)} \right) \right). \quad (3)$$

To eliminate y'_{AB} and F_{AB} , multiply (1) by

$$-a \left(\frac{1}{3} + \sum_{n=1}^s \frac{K_n}{n+3} \right)^*,$$

(2) by 1, (3) by $-1/a$ and add, then

$$\begin{aligned} EI_A y'_{BL} = EI_A \frac{y_B - y_A}{a} + a \left(\frac{1}{3} + \sum_{n=1}^s \frac{K_n}{n+3} \right) M_{BL} \\ + a \left(1 + \sum_{n=1}^s \frac{K_n}{n+1} - \frac{1}{2} - \sum_{n=1}^s \frac{K_n}{(n+1)(n+2)} - \frac{1}{3} - \sum_{n=1}^s \frac{K_n}{n+3} \right) M_{AR} \\ + \Sigma W_1 \left(\frac{(a-x_1)^2}{2} - \frac{a(a-x_1)}{3} - \frac{(a-x_1)^3}{2 \cdot 3a} + \sum_{n=1}^s K_n \left(\frac{a^{n+2}}{a^n(n+2)} - \frac{a^{n+1}x_1}{a^n(n+1)} \right. \right. \\ \left. \left. + \frac{x_1^{n+2}}{a^n(n+1)(n+2)} - \frac{a(a-x_1)}{n+3} - \frac{a^{n+3}-x_1^{n+3}}{a^{n+1}(n+2)(n+3)} + \frac{a^{n+1}x_1-x_1^{n+2}}{a^n(n+1)(n+2)} \right) \right), \end{aligned}$$

which simplifies to

$$\begin{aligned} EI_A y'_{BL} = EI_A \frac{y_B - y_A}{a} + a \left(\frac{1}{3} + \sum_{n=1}^s \frac{K_n}{n+3} \right) M_{BL} \\ + a \left(\frac{1}{2 \cdot 3} + \sum_{n=1}^s \frac{K_n}{(n+2)(n+3)} \right) M_{AR} \\ - \Sigma W_1 \left(\frac{x_1(a^2-x_1^2)}{2 \cdot 3a} + \sum_{n=1}^s \frac{K_n x_1 (a^{n+2}-x_1^{n+2})}{a^{n+1}(n+2)(n+3)} \right). \quad (4) \end{aligned}$$

§ 2. The distributed load over the bay may now be included. It may be regarded as the limit of $\sum w \delta X$, a series of concentrated loads over the bay, and will add to the right-hand side of (4), the term

$$- \int_0^a w dX \left(\frac{X(a^2-X^2)}{2 \cdot 3a} + \sum_{n=1}^s \frac{K_n X (a^{n+2}-X^{n+2})}{a^{n+1}(n+2)(n+3)} \right).$$

§ 3. The applied couples are next included. As the method used is believed to be new, it is developed in a general form.

Let the contribution of a downward force P at ξ be denoted by $P.f(\xi, \dots)$, where ξ is independent of the other variables. Then the contribution of an upward force P at $\xi + \delta\xi$ will be $-P.f(\xi + \delta\xi, \dots)$ or to the first order, $-P.f(\xi, \dots) - \delta\xi.P.\partial f(\xi, \dots)/\partial\xi$. Hence the contribution of the two forces together, forming $(\delta\xi \rightarrow 0)$ a counter-clockwise couple of moment $P.\delta\xi = M$ say at ξ , is $-P.\delta\xi.\partial f(\xi, \dots)/\partial\xi$, i. e., $-M\partial f(\xi, \dots)/\partial\xi$.

* This is
$$-a \left(\frac{1}{2} + \sum_{n=1}^s \frac{K_n}{n+2} - \frac{1}{2 \cdot 3} - \sum_{n=1}^s \frac{K_n}{(n+2)(n+3)} \right).$$

In equation (4) of § 1, the contribution to the right-hand side of load P_1 at ξ_1 is

$$-P_1 \left(\frac{\xi_1(a^2 - \xi_1^2)}{2.3a} + \sum_{n=1}^s \frac{K_n \xi_1(a^{n+2} - \xi_1^{n+2})}{a^{n+1}(n+2)(n+3)} \right)$$

hence the contribution of M_1 at ξ_1 is

$$M_1 \left(\frac{a^2 - 3\xi_1^2}{2.3a} + \sum_{n=1}^s \frac{K_n(a^{n+2} - (n+3)\xi_1^{n+2})}{a^{n+1}(n+2)(n+3)} \right),$$

and therefore

$$\Sigma M_1 \left(\frac{a^2 - 3\xi_1^2}{2.3a} + \sum_{n=1}^s \frac{K_n(a^{n+2} - (n+3)\xi_1^{n+2})}{a^{n+1}(n+2)(n+3)} \right)$$

is added to the right-hand side to allow for the applied couples ΣM_1 .

This method may be used in, and will shorten somewhat, the proofs of the first two forms of the Theorem of Three Moments previously referred to.

The method is also of general application in beam problems. Thus the deflexion at "x" for a simply supported beam, length l , due to load W at "a," is $Wx(l-a)(2al-a^2-x^2)/6EI$ (if $0 \leq x \leq a$) + $W(x-a)^3/6EI$ (if $a \leq x \leq l$). Differentiating partially with respect to "a" and replacing W by $-M$ gives the deflexion due to couple M at "a," viz.,

$$-Mx(2l^2 - 6al + 3a^2 + x^2)/6EI \text{ (if } 0 \leq x \leq a) + M(x-a)^2/2EI \text{ (if } a \leq x \leq l).$$

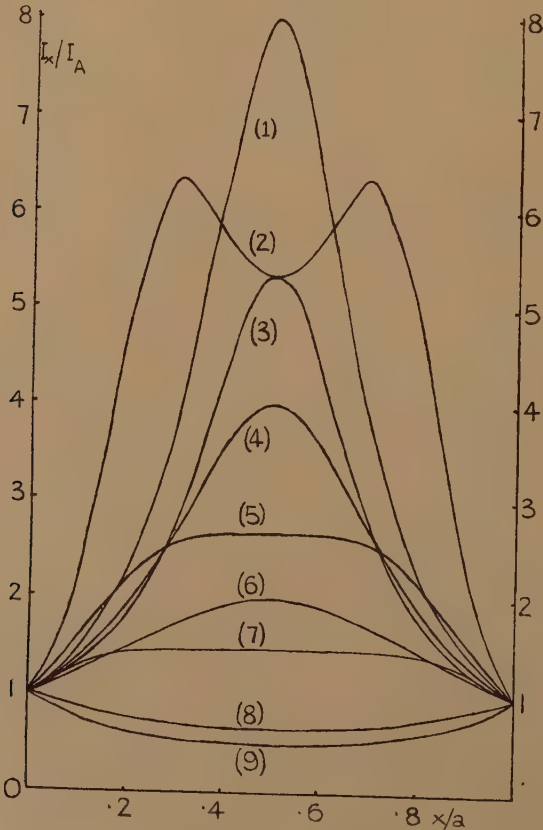
§ 4. A similar expression may be written down for $EI_C y'_{BR}$ from the bay CB. But $y'_{BL} + y'_{BR} = 0$, and substituting and rearranging, we get

$$\begin{aligned} & \frac{aM_{BL}}{EI_A} \left(\frac{1}{3} + \frac{K_1}{4} + \frac{K_2}{5} + \dots + \frac{K_s}{s+3} \right) \\ & + \frac{aM_{AR}}{EI_A} \left(\frac{1}{2.3} + \frac{K_1}{3.4} + \frac{K_2}{4.5} + \dots + \frac{K_s}{(s+2)(s+3)} \right) \\ & + \frac{cM_{BR}}{EI_C} \left(\frac{1}{3} + \frac{K_1'}{4} + \frac{K_2'}{5} + \dots + \frac{K_t'}{t+3} \right) \\ & + \frac{cM_{CL}}{EI_C} \left(\frac{1}{2.3} + \frac{K_1'}{3.4} + \frac{K_2'}{4.5} + \dots + \frac{K_t'}{(t+2)(t+3)} \right) \\ & = (y_A - y_B)/a + (y_C - y_B)/c \\ & + \Sigma \frac{W_1}{EI_A} \left(\frac{x_1(a^2 - x_1^2)}{2.3a} + \frac{K_1 x_1(a^3 - x_1^3)}{3.4a^2} + \frac{K_2 x_1(a^4 - x_1^4)}{4.5a^3} + \dots \right. \\ & \quad \left. + \frac{K_s x_1(a^{s+2} - x_1^{s+2})}{(s+2)(s+3)a^{s+1}} \right) \\ & + \Sigma \frac{W_1'}{EI_C} \left(\frac{z_1(c^2 - z_1^2)}{2.3c} + \frac{K_1' z_1(c^3 - z_1^3)}{3.4c^2} + \frac{K_2' z_1(c^4 - z_1^4)}{4.5c^3} + \dots \right. \\ & \quad \left. + \frac{K_t' z_1(c^{t+2} - z_1^{t+2})}{(t+2)(t+3)c^{t+1}} \right) \end{aligned}$$

$$+ \int_0^a \frac{w dX}{EI_A} \left(\frac{X(a^2 - X^2)}{2 \cdot 3a} + \frac{K_1 X(a^3 - X^3)}{3 \cdot 4a^2} + \frac{K_2 X(a^4 - X^4)}{4 \cdot 5a^3} + \dots + \frac{K_s X(a^{s+2} - X^{s+2})}{(s+2)(s+3)a^{s+1}} \right)$$

Fig. 2.
Symmetrical Examples.

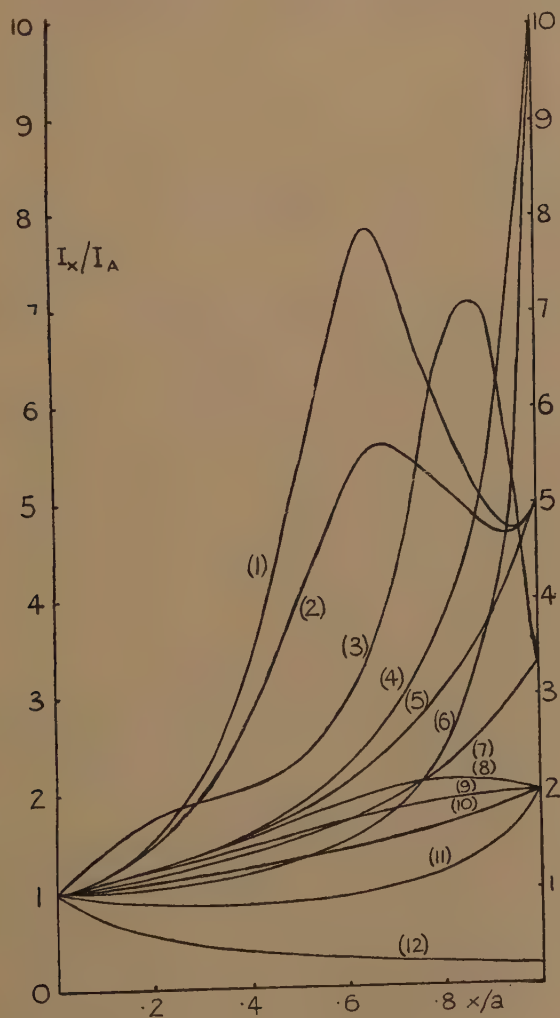
No.	1.	2.	3.	4.	5.	6.	7.	8.	9.
K ₁	-8	-3.5	-2	-3	-5	-2	-3	2	4
K ₂	27	3.5	-3	3	15	2	10	-2	-4
K ₃	-38	—	10	—	-20	—	-14	—	—
K ₄	19	—	-5	—	10	—	7	—	—



Form of $I/I_A = 1 / \left(1 + K_1 \frac{x}{a} + K_2 \frac{x^2}{a^2} + \dots \right)$.

Fig. 3.
Unsymmetrical Examples.

No...	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
$K_1 \dots$	-1.2	-1.2	-4.1	-1	-1	0	-0.7	-1	-1	-0.5	1	4
$K_2 \dots$	-4.6	-3.6	13.4	0.1	0.2	-1	—	0.1	0.5	—	-1.5	—
$K_3 \dots$	10	8	-20	—	—	0.1	—	0.4	—	—	—	—
$K_4 \dots$	-5	-4	10	—	—	—	—	—	—	—	—	—



$$\text{Form of } I_x/I_A = 1 / \left(1 + K_1 \frac{x}{a} + K_2 \frac{x^2}{a^2} + \dots \right).$$

$$\begin{aligned}
& + \int_0^c \frac{w' dZ}{EI_C} \left(\frac{Z(c^2 - Z^2)}{2 \cdot 3c} + \frac{K_1' Z(c^3 - Z^3)}{3 \cdot 4c^2} + \frac{K_2' Z(c^4 - Z^4)}{4 \cdot 5c^3} + \dots \right. \\
& \quad \left. + \frac{K_t' Z(c^{t+2} - Z^{t+2})}{(t+2)(t+3)c^{t+1}} \right) \\
& - \sum \frac{M_1}{EI_A} \left(\frac{a^2 - 3\xi_1^2}{2 \cdot 3a} + \frac{K_1(a^3 - 4\xi_1^3)}{3 \cdot 4a^2} + \frac{K_2(a^4 - 5\xi_1^4)}{4 \cdot 5a^3} + \dots \right. \\
& \quad \left. + \frac{K_s(a^{s+2} - (s+3)\xi_1^{s+2})}{(s+2)(s+3)a^{s+1}} \right) \\
& - \sum \frac{M_1'}{EI_C} \left(\frac{c^2 - 3\zeta^2}{2 \cdot 3c} + \frac{K_1'(c^3 - 4\zeta^3)}{3 \cdot 4c^2} + \frac{K_2'(c^4 - 5\zeta^4)}{4 \cdot 5c^3} + \dots \right. \\
& \quad \left. + \frac{K_t'(c^{t+2} - (t+3)\zeta_1^{t+2})}{(t+2)(t+3)c^{t+1}} \right),
\end{aligned}$$

where $M_{BR} = M_{BL} + m_B$.

A number of symmetrical and unsymmetrical examples of the variation in I_x/I_A for simple numerical values of "K"s up to K_4 , are given in fig. 2 and fig. 3.

§ 5. When the form of I_x is given throughout the bay, it will be necessary to find a polynomial of the form $1 + K_1x/a + K_2x^2/a^2 + \dots$ to approximate to I_A/I_x in the bay. The accuracy of the representation will, of course, usually increase with the degree of the polynomial, but the accuracy of M_{BL} etc. found from the Theorem will not necessarily so increase. Approximate values of K_1 , K_2 etc. may be calculated from the value of I_A/I_x at as many points as the number of K's to be included. These could be corrected by the method of least squares*, but this is laborious, and, it would appear from the example to be worked out in § 6, unnecessary for engineering accuracy.

If u_r denotes I_A/I_x at $r=x/a$, then

(A) for a polynomial of degree 2, $u_r = 1 + K_1r + K_2r^2$, K_1 and K_2 are found from $1 + \frac{1}{2}K_1 + \frac{1}{4}K_2 = u_{1/2}$, $1 + K_1 + K_2 = u_1$. These give

$$K_1 = -(3 - 4u_{1/2} + u_1), \quad K_2 = 2(1 - 2u_{1/2} + u_1).$$

(B) for a polynomial of degree 3, $u_r = 1 + K_1r + K_2r^2 + K_3r^3$, where

$$\begin{aligned}
K_1 &= -(11 - 18u_{1/3} + 9u_{2/3} - 2u_1)/2, \quad K_2 = 9(2 - 5u_{1/3} + 4u_{2/3} - u_1)/2, \\
K_3 &= -9(1 - 3u_{1/3} + 3u_{2/3} - u_1)/2.
\end{aligned}$$

(C) for a polynomial of degree 4, $u_r = 1 + K_1r + K_2r^2 + K_3r^3 + K_4r^4$, where

$$\begin{aligned}
K_1 &= -(25 - 48u_{1/4} + 36u_{1/2} - 16u_{3/4} + 3u_1)/3, \\
K_2 &= (70 - 208u_{1/4} + 228u_{1/2} - 112u_{3/4} + 22u_1)/3,
\end{aligned}$$

* See, e. g., Scarborough, 'Numerical Mathematical Analysis,' § 115.

$$K_3 = -16(5 - 18u_{1/4} + 24u_{1/2} - 14u_{3/4} + 3u_1)/3,$$

$$K_4 = 32(1 - 4u_{1/4} + 6u_{1/2} - 4u_{3/4} + u_1)/3.$$

When the bay is symmetrical about its middle, $u_r = u_{1-r}$, and in particular $u_1 = 1$. The labour of finding the K 's is reduced, and we get

(D) for degree 2, $u_r = 1 + K_1 r - K_1 r^2$, where $K_1 = -4(1 - u_{1/2})$.

(E) for degree 3, $u_r = 1 + K_1 r - K_1 r^2$, where $K_1 = -9(1 - u_{1/3})/2$. In this case K_3 vanishes, and we get a polynomial of degree 2 determined from $u_{1/3}$ instead of from $u_{1/2}$.

(F) for degree 4, $u_r = 1 + K_1 r + K_2 r^2 - 2(K_1 + K_2)r^3 + (K_1 + K_2)r^4$, where $K_1 = -(28 - 64u_{1/4} + 36u_{1/2})/3$, $K_2 = (92 - 320u_{1/4} + 228u_{1/2})/3$.

(G) for degree 5, $u_r = 1 + K_1 r + K_2 r^2 - 2(K_1 + K_2)r^3 + (K_1 + K_2)r^4$, where $K_1 = -25(5 - 9u_{1/2} + 4u_{1/4})/12$, $K_2 = 25(35 - 93u_{1/2} + 58u_{1/4})/24$.

K_5 is zero and we get a polynomial of degree 4 determined from $u_{1/2}$ and $u_{1/4}$ instead of from $u_{1/4}$ and $u_{1/2}$.

§ 6. To test the accuracy of the method, forms (D), (E), (F), and (G) have been applied to find M_B due to concentrated loads, each W , at the centre of each bay of a beam freely supported at A, B, and C at the same level, $AB = BC = a$, the beam being of uniform breadth, but increasing uniformly in depth to the centre of each bay, where it is double the depth at A, B, and C.

Thus $I_x = I_A(1 + 2x/a)^3$ ($0 \leq x \leq a/2$) and $u_r = (1 + 2x/a)^{-3} = (1 + 2r)^{-3}$, giving

$r = x/a = 0$	0.2	0.25	1/3	0.4	0.5
$u_r = 1$	0.36443	0.29630	0.21600	0.17147	0.125

Putting $M_{AR} = M_{CL} = 0$, $M_{BL} = M_{BR} = M_B$, $W_1 = W$, $x_1 = a/2$, the theorem reduces to

$$M_B(1/3 + K_1/4 + K_2/5 + K_3/6 + K_4/7 + \dots) = Wa(1/16 + 7K_1/192 + 3K_2/128 + 31K_3/1920 + 3K_4/256 + \dots).$$

(D) gives $K_1 = -3.5$, $K_2 = 3.5$, $M_B = .01693 Wa/.15833 = .1069 * Wa$,

(E) gives $K_1 = -3.528$, $K_2 = 3.528$, $M_B = .01656 Wa/.15693 = .1055 Wa$,

(F) gives $K_1 = -4.512$, $K_2 = 8.561$, $K_3 = -8.098$, $K_4 = 4.049$,

$$M_B = .01535 Wa/.14629 = .1049 Wa$$

(G) gives $K_1 = -5.012$, $K_2 = 11.514$, $K_3 = -13.002$, $K_4 = 6.501$,

$$M_B = .01588 Wa/.14484 = .1096 Wa.$$

* The last figure throughout is unreliable.

It should be noted that the numerator and denominator of M_B involve differences of relatively large quantities, and it is necessary to carry an extra figure in the K's.

The actual errors in the values of I_x/I_A from the polynomials, and also the percentage errors, at $r=x/a=.05$, $.95$, $.1$ and $.9$, etc., have been worked out and are exhibited in fig. 4*.

§7. We now require the true value of M_B for comparison.

If $0 \leq x \leq a/2$, and R be the reaction at A,

$$EI_x y'' = -Rx, \quad \text{where } I_x = I_A(1+2x/a)^3;$$

hence

$$EI_A y'' = -Rx(1+2x/a)^{-3}$$

and

$$2EI_A y''/a = R(1+2x/a)^{-3} - R(1+2x/a)^{-2},$$

$$4EI_A y'/a^2 = A - R(1+2x/a)^{-2}/2 + R(1+2x/a)^{-1},$$

$$8EI_A y/a^3 = A(1+2x/a) + B + R(1+2x/a)^{-1}/2$$

$$+ R \log(1+2x/a).$$

If $a/2 \leq x \leq a$,

$$EI_x y'' = -Rx + W(x-a/2), \quad \text{where } I_x = I_A(3-2x/a)^3;$$

hence

$$EI_A y'' = (-Rx + Wx - Wa/2)(3-2x/a)^{-3},$$

and

$$2EI_A y''/a = (R-W)(3-2x/a)^{-2} - (3R-2W)(3-2x/a)^{-3},$$

$$-4EI_A y'/a^2 = C - (R-W)(3-2x/a)^{-1} + (3R-2W)(3-2x/a)^{-2}/2,$$

$$8EI_A y/a^3 = C(3-2x/a) + D - (R-W) \log(3-2x/a) - (3R-2W)$$

$$(3-2x/a)^{-1}/2.$$

Now when $x=0$, $y=0$, hence

$$A+B+R/2=0, \quad \dots \dots \dots (i.)$$

when $x=a$, $y=0$; hence

$$C+D-(3R-2W)/2=0, \quad \dots \dots \dots (ii.)$$

where $x=a$, $y'=0$ by symmetry; hence

$$C-(R-W)+(3R-2W)/2=0. \quad \dots \dots \dots (iii.)$$

At $x=a/2$, y' is continuous; hence

$$A-R/8+R/2+C-(R-W)/2+(3R-2W)/8=0. \quad \dots \dots (iv.)$$

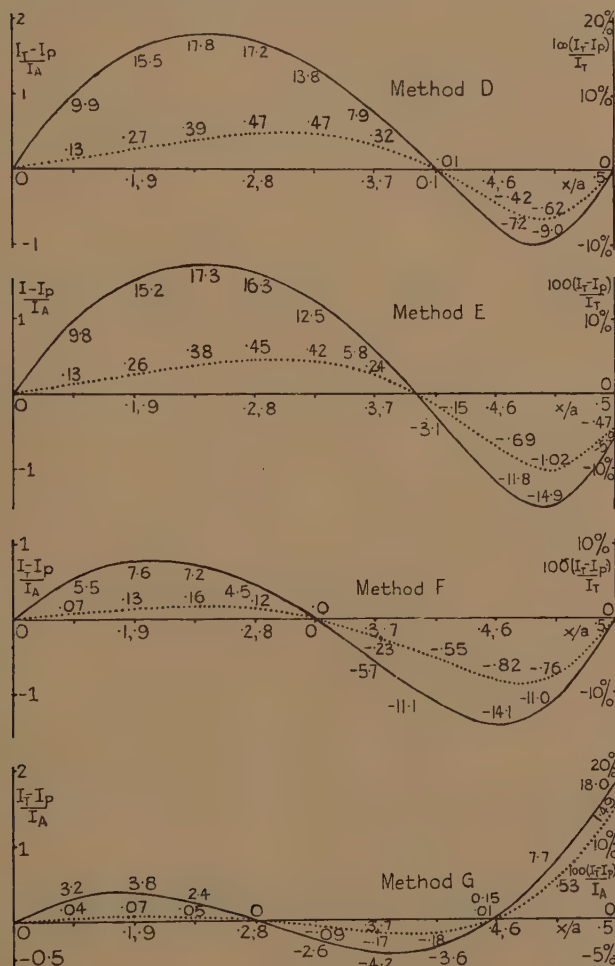
At $x=a/2$, y is continuous; hence

$$2A+B+R/4+R \log 2-2C-D+(R-W) \log 2+(3R-2W)/4=0. \quad (v.)$$

* The error at $x/a=.4$ in (G), being at a point from which the K's are calculated, should be zero. There is an actual percentage error of 0.15, due to neglect of practically a half-unit in the third place in rounding off both K_1 and K_2 and to the effects being additive.

Fig. 4.

- (i.) Actual error $(I_T - I_P)/I_A$ shown thus
 (ii.) Error per cent. $100(I_T - I_P)/I_T$ shown thus ————
 where I_T is the true value of I_x at "x."
 I_P is the value of I_x calculated from the polynomial.
 I_A is the value of I_x at A, where $x=0$.


 Errors in the values of I_x/I_A calculated from the polynomials.

Multiply (i.) by -1 , (ii.) by 1 , (iii.) by 2 , (iv.) by -1 , (v.) by 1 , and add, eliminating A, B, C, and D, and we find

$$R = W(4 \log 2 - 1)/(8 \log 2 - 1).$$

Hence $M_B = -Ra + Wa/2 = Wa/2(8 \log 2 - 1) = 1.1001 Wa.$

Thus all four values from the polynomials agree with the true value within 5 per cent.

§8. It is instructive also to examine the corresponding problem in which the additional depth is added in finite increments. Suppose this is achieved by adding "s" plates in each bay, the first from x_1 to x_{2s} , the next from x_2 to x_{2s-1} and so on, the last from x_s to x_{s+1} , where

$$0 < x_1 < x_2 \dots < x_s < a/2 < x_{s+1} \dots < x_{2s-1} < x_{2s} < a.$$

Let I_0, I_1 , etc. up to I_{2s} represent the second moment of the section over $0 \leq x \leq x_1$, $x_1 \leq x \leq x_2$, etc., up to $x_{2s} \leq x \leq a$ respectively. (In this example I_0 and I_{2s} will each equal I_A .)

Then we have

$$\begin{array}{lll} 0 \leq x \leq x_1 & x_1 \leq x \leq x_2 & \dots \text{ up to } x \leq n \leq a/2 \\ EI_0 y'' = -Rx & EI_1 y'' = -Rx & EI_s y'' = -Rx \\ EI_0 y' = A_0 - Rx^2/2 & EI_1 y' = A_1 - Rx^2/2 & EI_s y' = A_s - Rx^2/2 \\ EI_0 y = A_0 x + B_0 - Rx^3/6 & EI_1 y = A_1 x + B_1 - Rx^3/6 & EI_s y = A_s x + B_s - Rx^3/6. \end{array}$$

Now y' and y are continuous at $x = x_1$; hence

$$A_0/I_0 - Rx_1^2/2I_0 = A_1/I_1 - Rx_1^2/2I_1,$$

$$A_0 x_1/I_0 + B_0/I_0 - Rx_1^3/6I_0 = A_1 x_1/I_1 + B_1/I_1 - Rx_1^3/6I_1.$$

Multiply the first by $-x_1$ and add to the second, giving

$$B_0/I_0 + Rx_1^3/3I_0 = B_1/I_1 + Rx_1^3/3I_1. \quad \dots \quad (i.)$$

Similarly,

$$B_1/I_1 + Rx_2^3/3I_1 = B_2/I_2 + Rx_2^3/3I_2, \quad \dots \quad (ii.)$$

and so on to

$$B_{s-1}/I_{s-1} + Rx_s^3/3I_{s-1} = B_s/I_s + Rx_s^3/3I_s. \quad \dots \quad (s)$$

For $a/2 \leq x \leq x_{s+1}$, we have $EI_s y'' = -Rx + W(x - a/2)$, but by integrating by Macaulay's method, we make the constants of integration A_s and B_s apply throughout the interval x_s to x_{s+1} , giving

$$EI_s y' = A_s - Rx^2/2 + W(x - a/2)^2/2,$$

$$EI_s y = A_s x + B_s - Rx^3/6 + W(x - a/2)^3/6.$$

For $x_{s+1} \leq x \leq x_{s+2}$,

$$EI_{s+1} y'' = -Rx + W(x - a/2),$$

$$EI_{s+1} y' = A_{s+1} - Rx^2/2 + W(x - a/2)^2/2,$$

$$EI_{s+1} y = A_{s+1} x + B_{s+1} - Rx^3/6 + W(x - a/2)^3/6,$$

and for continuity of y' and of y at $x = x_{s+1}$, we get

$$\begin{aligned} A_s/I_s - Rx_{s+1}^2/2I_s + W(x_{s+1} - a/2)^2/2I_s \\ = A_{s+1}/I_{s+1} - Rx_{s+1}^2/2I_{s+1} + W(x_{s+1} - a/2)^2/2I_{s+1}, \end{aligned}$$

$$\begin{aligned} A_s x_{s+1}/I_s + B_s/I_s - R x_{s+1}^3/6I_s + W(x_{s+1}-a/2)^3/6I_s \\ = A_{s+1} x_{s+1}/I_{s+1} + B_{s+1}/I_{s+1} - R x_{s+1}^3/6I_{s+1} + W(x_{s+1}-a/2)^3/6I_{s+1}. \end{aligned}$$

Multiply the first by $-x_{s+1}$ and add to the second, giving

$$\begin{aligned} B_s/I_s + R x_{s+1}^3/3I_s - W(x_{s+1}-a/2)^2(2x_{s+1}+a/2)/6I_s \\ = B_{s+1}/I_{s+1} + R x_{s+1}^3/3I_{s+1} - W(x_{s+1}-a/2)^2(2x_{s+1}+a/2)/6I_{s+1}, \quad (s+1) \end{aligned}$$

and so on, to

$$\begin{aligned} B_{2s-1}/I_{2s-1} + R x_{2s}^3/3I_{2s-1} - W(x_{2s}-a/2)^2(2x_{2s}+a/2)/6I_{2s-1} \\ = B_{2s}/I_{2s} + R x_{2s}^3/3I_{2s} - W(x_{2s}-a/2)^2(2x_{2s}+a/2)/6I_{2s}. \quad (2s) \end{aligned}$$

Add corresponding sides of all the equations (i.) (ii.) ... (s) (s+1) ... (2s), we have

$$\begin{aligned} \frac{B_0}{I_0} + \frac{R}{3} \left(\frac{x_1^3}{I_0} + \frac{x_2^3}{I_1} + \dots + \frac{x_{2s}^3}{I_{2s-1}} \right) \\ - \frac{W}{6} \left(\frac{(x_{s+1}-a/2)^2(2x_{s+1}+a/2)}{I_s} + \dots + \frac{(x_{2s}-a/2)^2(2x_{2s}+a/2)}{I_{2s-1}} \right) \\ = \frac{B_{2s}}{I_{2s}} + \frac{R}{3} \left(\frac{x_1^3}{I_1} + \frac{x_2^3}{I_2} + \dots + \frac{x_{2s}^3}{I_{2s}} \right) \\ - \frac{W}{6} \left(\frac{(x_{s+1}-a/2)^2(2x_{s+1}+a/2)}{I_{s+1}} + \dots + \frac{(x_{2s}-a/2)^2(2x_{2s}+a/2)}{I_{2s}} \right). \end{aligned}$$

In this example,

$$x=0, y=0 \text{ gives } B_0=0,$$

$$x=a, y=0 \text{ gives } A_{2s}a + B_{2s} - Ra^3/6 + W(a/2)^3/6 = 0,$$

and by symmetry,

$$x=a, y'=0; \text{ hence } A_{2s} - Ra^2/2 + W(a/2)^2/2 = 0.$$

Multiply the second by $-a$, and add to the first, giving

$$B_{2s} + Ra^3/3 - 5Wa^3/48 = 0.$$

Also $I_0 = I_{2s} = I_A$; hence

$$\begin{aligned} \frac{R}{3} \left(\frac{x_1^3}{I_0} + \dots + \frac{x_{2s}^3}{I_{2s-1}} \right) \\ - \frac{W}{6} \left(\frac{(x_{s+1}-a/2)^2(2x_{s+1}+a/2)}{I_s} + \dots + \frac{(x_{2s}-a/2)^2(2x_{2s}+a/2)}{I_{2s-1}} \right) \\ = -\frac{Ra^3}{3I_A} + \frac{5Wa^3}{48I_A} + \frac{R}{3} \left(\frac{x_1^3}{I_1} + \dots + \frac{x_{2s}^3}{I_{2s}} \right) \\ - \frac{W}{6} \left(\frac{(x_{s+1}-a/2)^2(2x_{s+1}+a/2)}{I_{s+1}} + \dots + \frac{(x_{2s}-a/2)^2(2x_{2s}+a/2)}{I_{2s}} \right). \end{aligned}$$

Hence R^* and thence M_B since $M_B = -Ra + Wa/2$.

* R is a fraction whose numerator and denominator are differences of relatively large quantities, and it is necessary to carry an extra figure throughout.

Values of M_B have been worked out when the *additions* are of equal depth, placed symmetrically about the middle of each bay, and of lengths $(2s-1)a/2s$, $(2s-3)a/2s$, $a/2s$ respectively, so that $x_1=1/4s$, $x_2=3/4s$, $x_3=5/4s$, etc., as follows :—

One plate, viz., depth equal to depth at A, extending from $a/4$ to $3a/4$, gives	$M_B = \cdot 1 \quad Wa,$
Two plates, viz., depth of each half the depth at A, one from $a/8$ to $7a/8$, the other from $3a/8$ to $5a/8$, give	$M_B = \cdot 1001 \quad Wa,$
Three plates give	$M_B = \cdot 1043 \quad Wa,$
Four plates give	$M_B = \cdot 1062 \quad Wa,$
Five plates give	$M_B = \cdot 1076 \quad Wa,$
whilst the beam with uniformly increasing depth, previously worked out, represents the limit of the above as the number of plates tends to infinity, giving	$M_B = \cdot 1100 \quad Wa.$

In all cases, the polynomial approximations of §6. for a beam of depth increasing *uniformly* to the same central value, give values of M_B agreeing with these, for the "stepped" beam, within 10 per cent., and within 4 per cent. with four plates or more.

XXXI. *Theory of the Fractionation of Gaseous Mixtures by Diffusion.
The Characteristics of the Hertz-Mercury-Vapour-Pump.*

By E. BLUMENTHAL *.

[Received December 7, 1938.]

§1. *Introduction.*

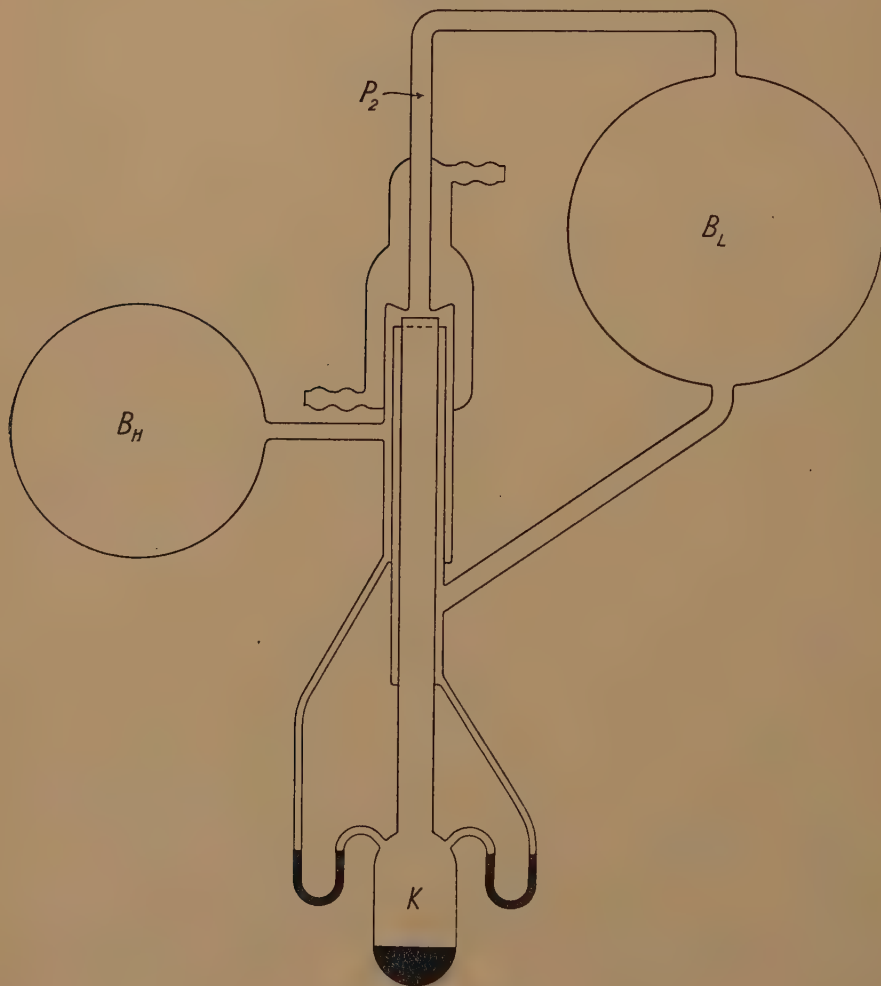
THE possibility of fractionating gaseous mixtures, even mixtures of isotopic gases, by diffusion into a moving gaseous medium has been proved theoretically and experimentally by Hertz⁽¹⁾. In his original experimental arrangement the separation was obtained by a single diffusion process. This gave good results for a mixture of He and Ne, but for the separation of isotopic gases the method only became practicable with the construction of a diffusion unit, which can be worked "in series" (figs. 1 *a*, 1 *b*)⁽²⁾. The unit is beautifully compact: the mercury vapour, which serves as diffusion medium, is used at the same time to pump the isotopic mixture from one "diffuser" to the other. The properties of such a "diffusion pump" have been studied in detail experimentally and theoretically by H. Barwich⁽³⁾; they have been discussed further by D. MacGillavry⁽⁴⁾, and by R. Sherr⁽⁵⁾, to whom several modifications of the original apparatus are due.

In his theoretical analysis, Barwich finds that the fractionation obtained with such a diffuser should increase exponentially with the velocity v of the stream of mercury vapour. Experimentally he finds an increase of the fractionation q with v at low values of v , and a decrease at higher values (cf. figs. 2 *a*, 2 *b*). Between the initial increase and final decrease of q either a single maximum or two maxima and a minimum, depending on the introduction of a capillary resistance into the tube P_1P_2 at P_2 (cf. fig. 1 *b*), are observed. These discrepancies between theory and experiment have been ascribed by the above workers to various irregularities in the conditions under which the diffusion takes place, *e. g.*, turbulence of the mercury vapour stream. As the possibility of improving upon Hertz's model of the diffuser depends on the correct interpretation of these discrepancies, a new analysis of the conditions governing the process of fractionation is undertaken.

* Communicated by Prof. M. Polanyi.

§ 2. *The Process of Fractionation.*

Fig. 1 *a* represents a single diffusion unit, fitted with the end volumes B_L and B_H , in which, respectively, the light and heavy fractions of the

Fig. 1 *a*.

gaseous mixture filling the whole apparatus are accumulated. Fig. 1 *b* shows only the part where the actual diffusion takes place. Both figures are drawn to scale, only the distance BP in fig. 1 *b* is exaggerated three-fold.

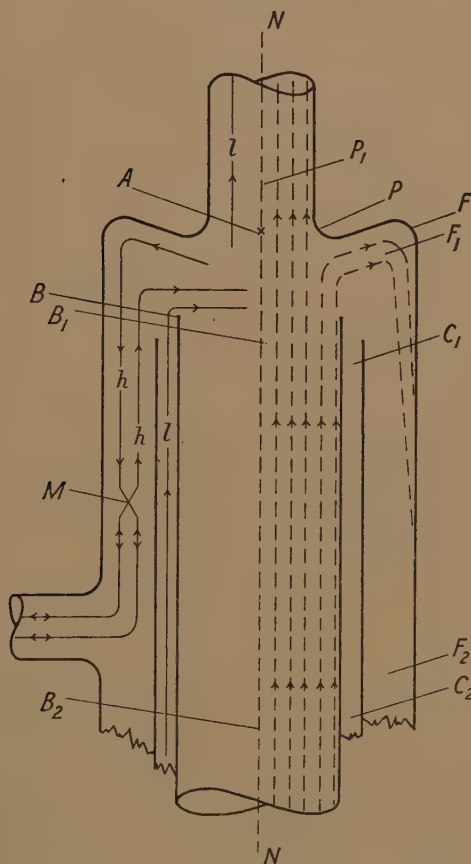
Mercury is boiled in K, and the stream of mercury vapour produced

flows up the tube B_2B_1 with the velocity v cm./sec. If the rate of heating is denoted by W , then in rough approximation

$$v \propto W - W_0, \dots \dots \dots (1)$$

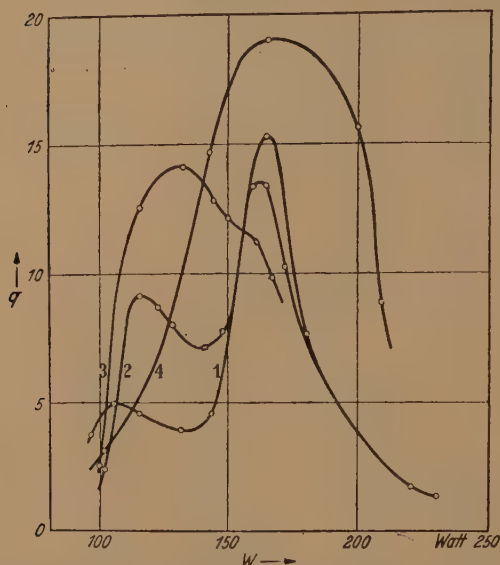
where W_0 represents heat losses. The stream of mercury vapour and the

Fig. 1 b.



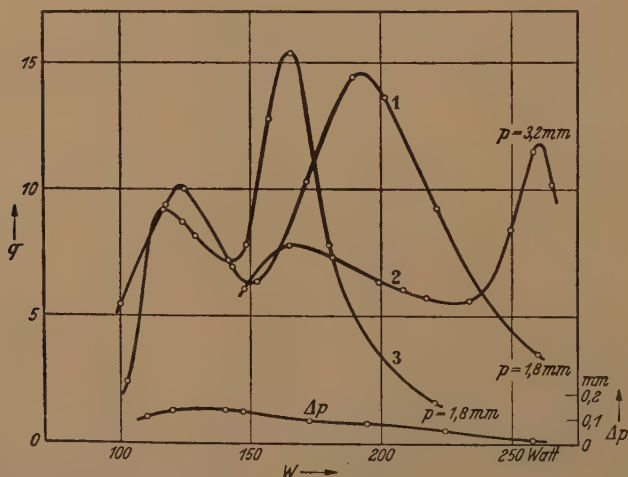
gaseous mixture come into contact above the edge B , the mercury vapour forming a jet, across which diffusion of the mixture of gases takes place. The edge P splits the stream of mercury vapour into a central portion, going up the tube P_1P_2 , and an outer portion, streaming along the wall PF and down the annulus F_1F_2 . The streamlines of the mercury vapour are shown schematically in fig. 1 b. As the lighter molecules of the gaseous mixture have a bigger chance to reach the central part

Fig. 2 a.



Experiments with pump 1. Pressure = 1.8 mm. Hg. Curve 1: duration of experiments, 20 mins.; apparatus as sketched in fig. 1 a. Curve 2: duration, 20 mins.; capillary (5 cm. long, 0.3 cm. wide) inserted into tube P_1 at P_2 . Curve 3: duration, 40 mins.; capillary 3 cm. long, 0.1 cm. wide. Curve 4: duration, 60 mins., connexion between B_L and annulus C_2C_1 interrupted at C_2 (cf. fig. 1 b).

Fig. 2 b.



Duration of experiments, 15 mins. Δp = difference of pressures between light and heavy sides. Curves 1, 2 obtained with pump 2, curve 3 with pump 1, curve for Δp with pump 2.

of the stream of mercury vapour in the interval of time t_0 , during which the stream travels from B to P, the part of the diffusate which is carried up the tube P_1P_2 is richer in the light component than the diffusing mixture. Conversely the diffusate carried into the annulus F_1F_2 by the outer portion of the mercury vapour stream is richer in the heavy component. The light fraction of the diffusate is pumped through the tube P_1P_2 into the end-volume B_L , mixes with the gas contained therein, and an equal amount of gas returns to the diffuser through the annulus C_2C_1 . The path of the light fraction is shown schematically by the line l, l in fig. 1 *b*. The heavy fraction of the diffusate is carried into the annulus F_1F_2 by the stream of mercury vapour. Equilibration of concentrations between the heavy fraction of the diffusate and the gas contained in the end-volume B_H takes place near that region of the annulus F_1F_2 , denoted by M in fig. 1 *b*. This equilibration is brought about partly by diffusion, partly by irregular convection currents set up by the condensing mercury,—condensation not being a smooth process, but proceeding *via* the stages: supersaturation, formation of nuclei, rapid condensation and production of partial vacua. Whether this equilibration is complete or only partial depends on a number of factors, but the existence of these irregular convection currents certainly makes it approach complete equilibration much closer than if it proceeded by diffusion alone. After equilibration, complete or partial, has taken place, an amount of the resulting mixture equal to that of the heavy fraction of diffusate returns to the region above the edge BB, to re-diffuse into the mercury vapour stream; thus the distribution of pressures in the apparatus is kept constant. In fig. 1 *b* the path of the heavy fraction is indicated by the line h, h .

A stationary distribution of the components of the mixture of gases in the apparatus, *i. e.*, “equilibrium of compositions,” is reached, when the light fraction of the diffusate is of the same composition as the gas in B_L , and the heavy fraction as that in B_H .

Notation.—Before proceeding to a quantitative study of the process of fractionation a system of notation must be introduced. The partial pressures of the light, heavy component of the gaseous mixture at any position in the apparatus are denoted by l, h respectively; for convenience l, h are measured in gm.mols./c.c. The composition of the mixture is defined as the ratio of the partial pressure of the heavy to that of the light component, and denoted by ρ . Hence $\rho = h/l$. Rates of flow of the various gases in the apparatus are denoted by capital letters: L denoting the rate of flow of the light component at any position in the apparatus, H that of the heavy component, G that of the mixture of gases. The unit is: gm.mols./sec.

The position in the apparatus, at which the various partial pressures,

compositions, rates of flow, are measured, is denoted by a subscript; the time λ , at which the measurement is taken, by a second subscript. The following positional subscripts are used: Subscripts l, h, m , denote measurements on the diffusate, taken in the stream of mercury vapour at the height of the edge P. Subscripts l, h , refer to the light, heavy fraction of the diffusate respectively, m to the total diffusate. *E. g.*, l_h is the *average* partial pressure of the light component of the diffusate in the outer part of the stream of mercury vapour, measured at the height PP; ρ_l is the *average* composition of the diffusate in the central part of the stream of mercury vapour, going up the tube P_1P_2 , measured at the height PP; G_l is the total flow of diffusate up the tube P_1P_2 . Subscripts L, H, refer to the "light" and "heavy" sides of the apparatus respectively, the "light" side comprising the "light" end-volume B_L , the annulus C_2C_1 , and the tubes connecting B_L to C_2C_1 and P_1P_2 , and the "heavy" side the "heavy" end-volume B_H , the annulus F_2F_1 up to the region M, and the tube connecting both. Subscript s refers to the "source of diffusion." The "source of diffusion" is a concept introduced for convenience of expression; it is imagined as a short, thin annulus, adjacent to the interface of the jet of mercury vapour and the mixture of gases, and extending from the height of the edge B to that of P. All the gas diffusing into the stream of mercury vapour passes through this annulus, and can therefore be said to diffuse from it. Subscript o refers to the gas flowing into the source of diffusion.

Hence we have by definition:

$$L_m = L_l + L_h, H_m = H_l + H_h, G_m = G_l + G_h. \quad (2)$$

In terms of this notation the compositions of the light, heavy fractions are given by H_l/L_l , H_h/L_h respectively. Hence the condition for equilibrium of compositions, derived above, reads:

$$\rho_{L\infty} = \frac{H_{l\infty}}{L_{l\infty}}, \quad (3)$$

$$\rho_{H\infty} = \frac{H_{h\infty}}{L_{h\infty}}. \quad (4)$$

Coefficients of Fractionation.—The coefficient of fractionation, q , of a diffusion unit at a fixed value of v is defined as the ratio of the compositions of the gases in the heavy and light sides, when equilibrium of compositions has been reached, *i. e.*, at $t = \infty$. Hence

$$q = \frac{\rho_{H\infty}}{\rho_{L\infty}}, \quad (5)$$

q is independent of the relative dimensions of the light and heavy side and the initial composition of the mixture of gases. To describe the

fractionation by diffusion into the stream of mercury vapour we introduce the "coefficients of fractionation" for the light fraction, q_l , and "the coefficient of fractionation for the heavy fraction," q_h , so that

$$q_l = \frac{L_l}{H_l} \bigg/ \frac{L_o}{H_o} \dots \dots \dots (6)$$

$$q_h = \frac{H_h}{L_h} \bigg/ \frac{H_o}{L_o} \dots \dots \dots (7)$$

The relation between q and q_l , q_h is found by substitution from equations 3, 4, 6, 7 in equation (5); this gives directly

$$q = q_l q_h \dots \dots \dots (8)$$

Hence q can be obtained by evaluation of q_l and q_h .

§3. Theoretical Estimation of the Coefficient of Fractionation.

(a) Model of the Process of Diffusion.

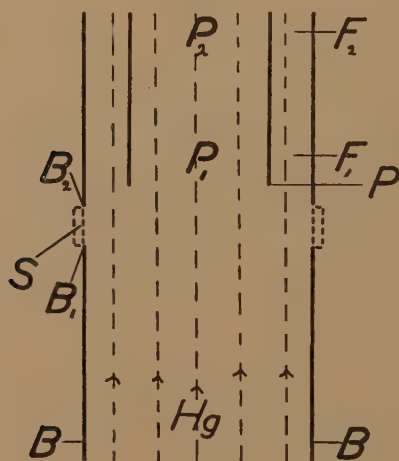
In order to derive theoretical expressions for q , Hertz⁽¹⁾ and McGillavry⁽⁴⁾ have made two assumptions: (1) D , the coefficient of interdiffusion, is independent of the ratio in which the interdiffusing gases are mixed, (2) the rates of diffusion of two gases diffusing from the same source into any medium of diffusion are independent of each other. McGillavry⁽⁴⁾ has shown from Jeans's theoretical work that assumption (1) only entails a very slight error. According to Trautz and co-workers⁽⁶⁾, assumption (1) is in practical agreement with experimental results. The validity of assumption (2) for the case of the Hertz pump is proved by evidence obtained by Von Hartel, Meer and Polanyi⁽⁷⁾ in experiments on the diffusion of sodium vapour through mixtures of gases. These authors showed that the diffusion coefficient of sodium vapour into argon, or nitrogen, is not altered by admixture of up to 70 per cent. hydrogen to the argon, or nitrogen, whereas even 1 per cent. of argon, or nitrogen, in hydrogen depresses the diffusion coefficient of sodium vapour markedly from the value for pure hydrogen. It appears that collisions with the heavy constituent of the diffusing mixture are far more important for the progress of diffusion than those with the light constituents. Hence assumption (2) holds in all cases in which the gaseous mixture to be fractionated in the Hertz diffuser is light compared with mercury vapour.

To construct a model, suitable for calculations, of the conditions under which the separation by diffusion takes place, three simplifying assumptions are introduced: (a) The pressure, p , and the temperature, T , in the diffuser are assumed to be constant, independent of v . Experimentally slight variations of p with v are found (*cf.* fig. 2 *b*), but no significant connexion between Δp and q (*loc. cit.*⁽³⁾, p. 178). (b) The stream of mercury vapour is split up by the edge PP into a central and an outer

part in a ratio independent of v (*cf.* paragraph 5). (c) In the direction of flow the stream of mercury vapour is assumed to be infinitely long, and the pumping action is neglected. This assumption is most likely to lead to errors: (1) if v is very low, and the stream of mercury vapour does not considerably extend past the source of diffusion, (2) if the capillary resistance in the path of the light fraction is big. But the errors are found to be of no practical importance.

In contrast to models of the diffusion process previously suggested, we find it necessary that our model should satisfy the following two conditions: (1) It must lend itself to an evaluation of the amounts of diffusate in both fractions; hence, as in the actual case, the concentration

Fig. 3.



of diffusate in the source of diffusion must be finite. (2) It must allow the variation with v of the amount of gas diffusing per unit time to be calculated. This necessitates that, as in the actual case, the source of diffusion is extended in the direction of flow of the mercury vapour stream.

These conditions are satisfied by the model sketched in fig. 3. The stream of mercury vapour flows up the tube BB and comes into contact with the mixture of gases outside the tube along a short, but finite length B_1B_2 , where the walls of the tube are cut away. The source of diffusion, imagined as a short thin annulus, is indicated by the areas S .

We consider the diffusion of a single gas l and choose cylindrical co-ordinates so that the z -axis coincides with the axis of the tube B , and the direction of flow of the stream of mercury vapour with the direction of the

positive z -axis. Then, according to Hertz⁽¹⁾, the differential equation obeyed by the diffusing gas is

$$\nabla^2 l = \frac{v}{D} \frac{\partial l}{\partial z}, \quad \dots \dots \dots (9)$$

which in cylindrical coordinates takes the form (*cf. loc. cit.*⁽⁴⁾)

$$\frac{\partial^2 l}{\partial r^2} + \frac{1}{r} \frac{\partial l}{\partial r} + \frac{\partial^2 l}{\partial z^2} = \frac{v}{D} \frac{\partial l}{\partial z}. \quad \dots \dots \dots (10)$$

A solution of equation (10) for a set of boundary conditions, including the two given above, has been worked out by F. Murray⁽⁸⁾, but it is far too cumbersome to be of practical use. Mainly responsible for the complexity of the solution is condition (2), which we wish to retain unaltered.

The problem, however, can be simplified further: (*d*) Evaluation of equation (10) with the proper boundary conditions and the two values of D for the diffusion of the two components of the gaseous mixture into the stream of mercury vapour allows an exact calculation of q_l , q_h , and hence q , at any value of v . But this is more than we need, as for our purpose we are far less interested in the absolute values of q , than in the possibility of finding a theoretical explanation of the stationary values of q found experimentally on varying v . As these stationary values are very marked, all we require is an approximate theory of the change of q with v .

The reason for using a moving diffusion medium in a fractionating device can be thought of as follows: diffusion into a medium at rest leads to fractionation, as long as the steady state is not reached. This is demonstrated by the differential equation of diffusion into a medium at rest,

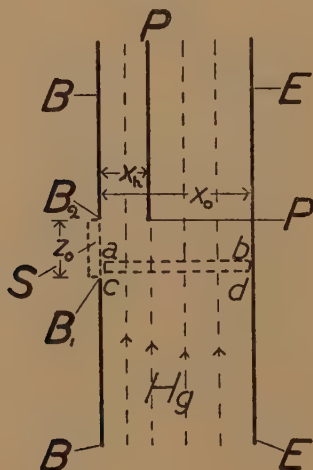
$$\frac{\partial l}{\partial t} = D \nabla^2 l, \quad \dots \dots \dots (11)$$

in the solution of which D occurs, provided that $\frac{\partial l}{\partial t} \neq 0$. So, if after a short time the diffusion medium is taken out of the reach of the source of diffusion, and the parts which have been farther or nearer to the source of diffusion are separated from one another, they will contain a lighter or heavier fraction of the diffusate respectively. This is the purpose of having a moving diffusion medium. Fig. 4 illustrates this concept; the diffusion medium streams between infinite parallel plates B, P, E—introduced as a further simplification instead of the coaxial cylinders of fig. 3—in the direction of the positive z -axis. The diffusing gases are to the left of plate B, and come into contact with the diffusion medium along a strip cut out of B parallel to the y -axis, the x -axis being perpendicular to B. The source of diffusion is indicated by the area S.

We consider the element of diffusion medium enclosed between the neighbouring parallel planes ab and cd , moving with the velocity of the diffusion medium. While it travels past the source of diffusion gas diffuses into it, as into a cylinder containing a diffusion medium at rest. After it has passed the source of diffusion it is split up by the plane P into two parts, the part between the planes P and E taking with it the light, and the part between the planes P and B the heavy fraction of the diffusate.

This model of the diffusion process satisfies the two conditions stated above. The only deviation from the actual process of diffusion—apart from the points stated in assumptions a , b , c —is that the diffusion parallel to the z -axis is neglected. That this diffusion is unimportant for our

Fig. 4.



problem cannot be proved except by solution of equation (10), but it is made rather probable by the way the model is arrived at.

(b) *Partial Pressures and Rates of Flow of Diffusate.*

To find the distribution of diffusate in the diffusion medium at the height of the edge P according to our model, the partial pressure l of diffusate at any point of a cylinder of length x_0 , into which diffusion from a source of constant pressure l_s has taken place for the time $t_0 = \frac{z_0}{v}$, must be calculated (cf. fig. 4). This problem has been solved by E. L. Lederer⁽⁹⁾. The differential equation is

$$\frac{\partial l}{\partial t} = D \frac{\partial^2 l}{\partial x^2} \quad \dots \dots \dots (12)$$

The boundary conditions are :

$$\begin{aligned} 0 \leq t \leq \infty, \quad x=0, \quad l=l_s; \\ x=x_0 \quad \frac{\partial l}{\partial x}=0. \end{aligned}$$

The solution is

$$l=l_s \left\{ 1 - \text{SIN} \left(\frac{x}{x_0}, ft_0 \right) \right\}, \quad . \quad . \quad . \quad . \quad . \quad (13)$$

where

$$f = \frac{\pi^2 D}{4x_0^2},$$

and

$$\text{SIN} \left(\frac{x}{x_0}, ft \right) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin \frac{2n+1}{2} \frac{\pi x}{x_0} e^{-(2n+1)^2 ft}.$$

To calculate the rates of flow of diffusate L_m , L_l , and L_h Hertz's relation⁽¹⁾ connecting the flow and the partial pressure of the diffusate at any point of the diffusion medium

$$\bar{L} \propto \bar{v} l - D \text{ grad } l \quad . \quad . \quad . \quad . \quad . \quad (14)$$

(in our notation) is used. For the flow along the z -axis this reduces to

$$L \propto v l \quad . \quad . \quad . \quad . \quad . \quad (15)$$

in our model. The ratio of the radii of the tubes P_1P_2 and B_1B_2 in Hertz's diffuser is $\sim \frac{2}{3}$, hence we let the distance x_h between the planes P and E be $\frac{1}{3}$ of the distance x_0 between the planes E and B. Substituting in equation (13) the definitions of l_m and l_l , we obtain :

$$x_0 l_m = l_s \int_0^{x_0} \left\{ 1 - \text{SIN} \left(\frac{x}{x_0}, ft_0 \right) \right\} dx \quad . \quad . \quad . \quad . \quad . \quad (16)$$

and

$$\left(x_0 - \frac{x_0}{3} \right) l_l = l_s \int_{\frac{x_0}{3}}^{x_0} \left\{ 1 - \text{SIN} \left(\frac{x}{x_0}, ft_0 \right) \right\} dx. \quad . \quad . \quad . \quad (17)$$

The solution of equation (16) is⁽⁹⁾

$$l_m = l_s \{ 1 - \text{INT} (ft_0) \}, \quad . \quad . \quad . \quad . \quad . \quad (18)$$

where

$$\text{INT} (ft) = \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \cos \frac{(2n+1)\pi}{6} e^{-(2n+1)^2 ft}.$$

The solution of equation (17) is

$$l_l = l_s \left\{ 1 - \frac{12}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \cos \frac{(2n+1)\pi}{6} e^{-(2n+1)^2 ft_0} \right\} \quad . \quad . \quad . \quad (19)$$

Substituting, in equation (15), we find for L_m and L_l ,

$$L_m = \lambda v x_0 l_s \left\{ 1 - \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} e^{-(2n+1)^2 \pi^2 l_s} \right\}, \quad (20)$$

$$L_l = \frac{2}{3} v x_0 l_s \left\{ 1 - \frac{12}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \cos \frac{(2n+1)\pi}{6} e^{-(2n+1)^2 \pi^2 l_s} \right\}, \quad (21)$$

λ being the length of the source of diffusion along the y -axis. l_h is found by the relation

$$\left(x_0 - \frac{x_0}{3}\right) l_l + \frac{x_0}{3} l_h = x_0 l_m,$$

which follows from the definition, and L_h by equation (2).

For the evaluation of equations (18) to (21) substitution of numerical values for D , z_0 , etc., is necessary. We shall take those appropriate to the enrichment of H_2O^{18} in water vapour in which we are more specially interested. $p=2.5$ mm. Hg. According to Knudsen's formula the vapour pressure of mercury equals 2.5 mm. Hg. at about 146°C . Hence $T=419^\circ$ abs. In Hertz's diffuser the vertical distance $\text{BP}=3$ mm. Only part of this corresponds to z_0 ; hence we take $z_0=2$ mm. x_0 is taken as 9 mm., whereas the radius of tube B is 10 mm. No experimental determination of D for the interdiffusion of water and mercury vapour is known to us. Hence D is calculated, using a set of values of $D(\text{H}_2\text{O}/\text{CO}_2)$ at temperatures up to 92.4°C .⁽¹⁰⁾ These follow the relation $D \propto T^2$ fairly closely, which is also given by an equation of Maxwell's⁽¹¹⁾,

$$D_{12} = \frac{1}{2hm_1m_2(v_1+v_2)A_1} \sqrt{\frac{m_1+m_2}{K}}, \quad (22)$$

where $A_1 = \text{const}$, $h \propto \frac{1}{T}$, $m_{1,2}$ = molecular mass, and $v_{1,2}$ = partial pressure of the interdiffusing gases 1, 2 respectively. This equation holds rigorously for any values of $\frac{v_1}{v_2}$ and $\frac{m_1}{m_2}$. Provided the law of force between the molecules 1, 2 is

$$F = -m_1m_2 \frac{K}{r^5}.$$

Using equation (22) we find for 146°C . and 2.5 mm. Hg :

$$D(\text{H}_2\text{O}/\text{Hg}) = 39 \text{ cm.}^2/\text{sec.} \approx 40 \text{ cm.}^2/\text{sec.}$$

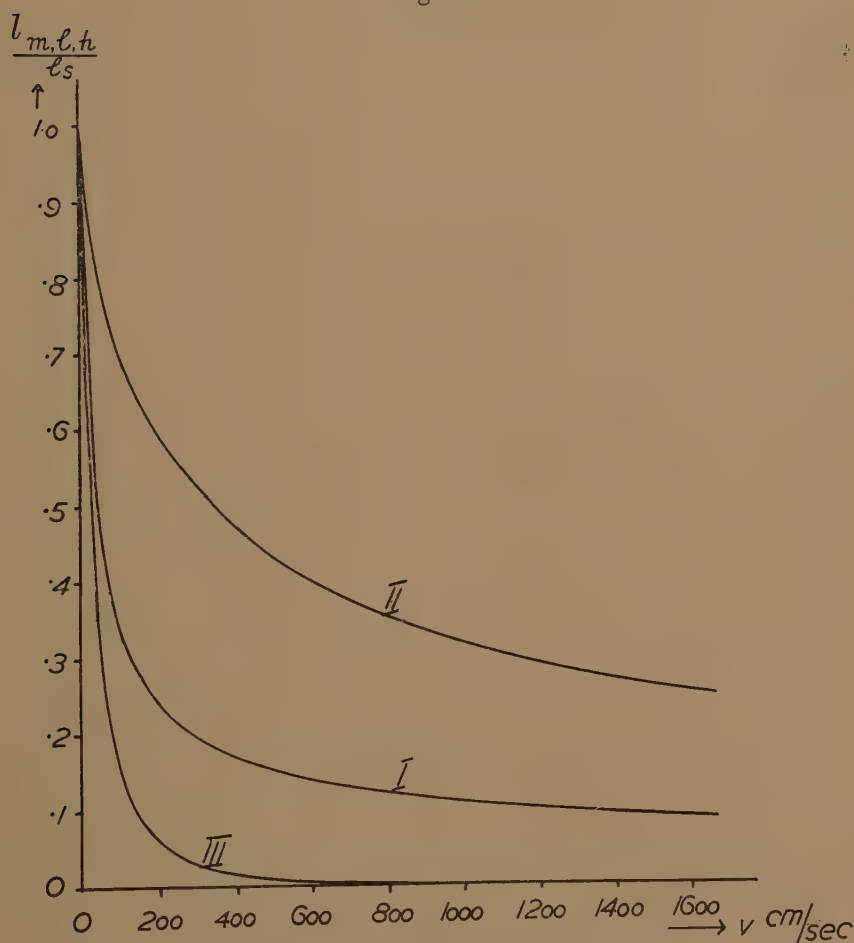
The exponential series of equations (18)–(21) are evaluated to six significant figures.

The results are shown in figs. 5 and 6. The scale along the y -axis relates to $\frac{l_h, l, m}{l_s}$ in fig. 5, and to $\frac{L_h, l, m}{\lambda l_s x_0}$ in fig. 6.

Corollaries.

(1) The variation of the amount of diffusate in the light fraction, L_l , with v explains the behaviour of a light glass sphere (0.1 gm.) put into the part of the light fraction, which is used by Barwich as indicator (*loc. cit.*⁽³⁾,

Fig. 5.



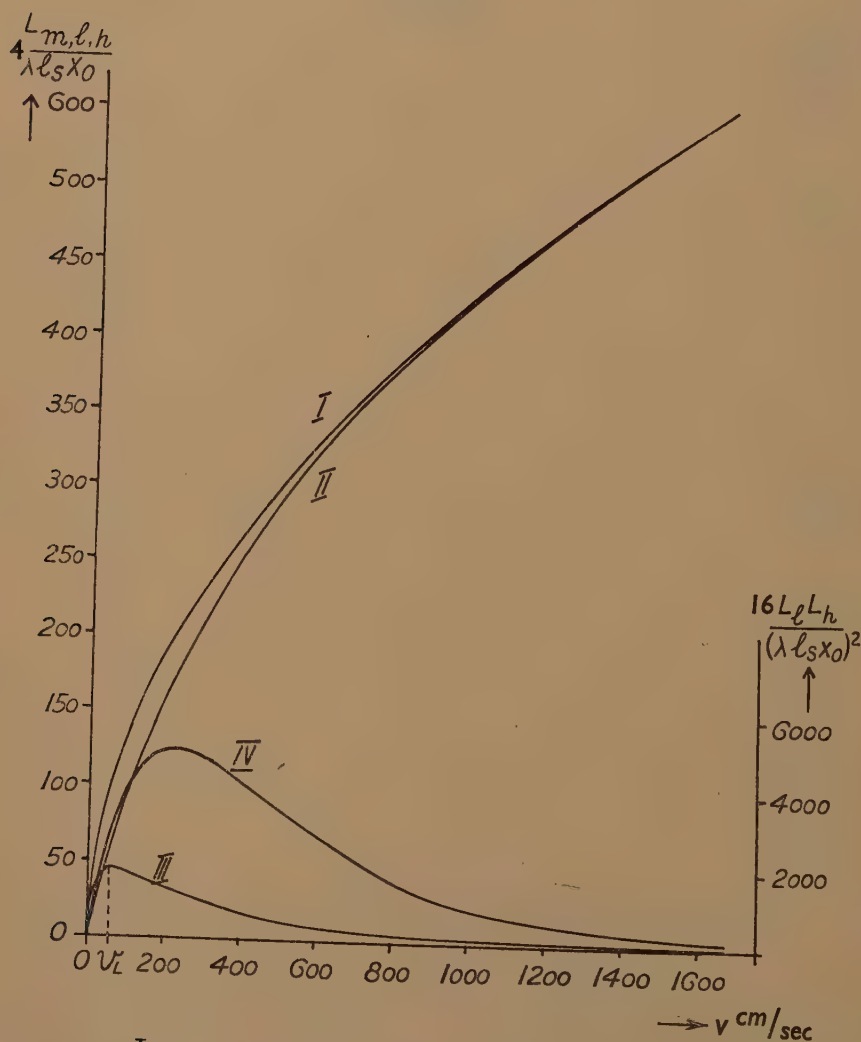
Curve I. represents l_m/l_s , curve II l_l/l_s , curve III. l_h/l_s .

p. 174). It is lifted up at low rates of heating, and sinks at higher rates (*cf.* paragraph 5).

(2) The difference in pressure Δp between the light and heavy sides (*cf.* fig. 2 b) must be due to gas being pumped into the heavy side, while

equilibration of pressures between the two sides is hindered by the capillary (5 cm. long, 3 mm. wide) in P_1P_2 and the streaming pressure of the stream of gas in C_1C_2 . Hence Δp will vary with v as $(L_l L_h)$; actually the curves for Δp and $(L_l L_h)$ agree in shape (cf. paragraph 5).

Fig. 6.



$\frac{L_m}{\lambda l_s x_0}$ represented by curve I., $\frac{L_h}{\lambda l_s x_0}$ by curve II.,

$\frac{L_l}{\lambda l_s x_0}$ by curve III., $\frac{L_l L_h}{(\lambda l_s x_0)^2}$ by curve IV.

(c) *Evaluation of the Coefficients of Fractionation.*

Under assumption (2), paragraph 3 (a), evaluation of equations (18) and (19) for the two values of D , D_l , and D_h , leads to the values of ρ_l/ρ_s and ρ_h/ρ_s . To derive q_l and q_h from these magnitudes it is necessary to know the relation connecting ρ_s and ρ_o . Though the composition of the gaseous mixture in the source of diffusion is brought to the value ρ_s by the inflow of gas of composition ρ_o , ρ_s need not equal ρ_o , as the continuous loss of gas from the source by diffusion into the stream of mercury vapour does not allow equilibration of ρ_s and ρ_o . But ρ_s is fixed by the condition that as much of either component as flows into the source of diffusion in unit time must leave it by diffusion into the stream of mercury. Hence by equation (20)

$$\frac{L_o}{H_o} = \frac{L_m}{H_m} = \frac{l_s \{1 - \text{INT} (f_l t_o)\}}{h_s \{1 - \text{INT} (f_h t_o)\}} \dots \dots \dots (22 a)$$

As

$$\frac{H_o}{L_o} = \rho_o \quad \text{and} \quad \frac{h_s}{l_s} = \rho_s,$$

$$\frac{\rho_s}{\rho_o} = \frac{1 - \text{INT} (f_l t_o)}{1 - \text{INT} (f_h t_o)} \dots \dots \dots (23)$$

In paragraph (3 b) $D_l \approx 40$ cm.²/sec. has been found. To calculate D_l/D_h we can either use equation (22), which leads to the value 1.09, or another equation of Maxwell's (*loc. cit.*⁽¹¹⁾, p. 316),

$$D_{12} = \frac{2}{3\pi(\nu_1 + \nu_2)s_{12}^2} \sqrt{\frac{1}{\pi h} \left(\frac{1}{m_1} + \frac{1}{m_2} \right)}, \dots \dots \dots (24)$$

derived on the assumption that molecules behave like elastic spheres, s_{12} being the sum of the collision radii of the two kinds of molecules. Assuming the collision radii of H_2O^{16} and H_2O^{18} to be identical, equation (24) leads to $D_l/D_h = 1.03$. Hence we assume the real value of D_l/D_h to lie between 1.03 and 1.09. For purposes of calculation we take $D_l/D_h = 1.1$. The results are shown in fig. 7.

Comparison of figs. 6 and 7 shows that the decrease of L_l (and H_l) with v is much quicker than the increase of q . This suggests that the velocity of separation may decrease rapidly with v .

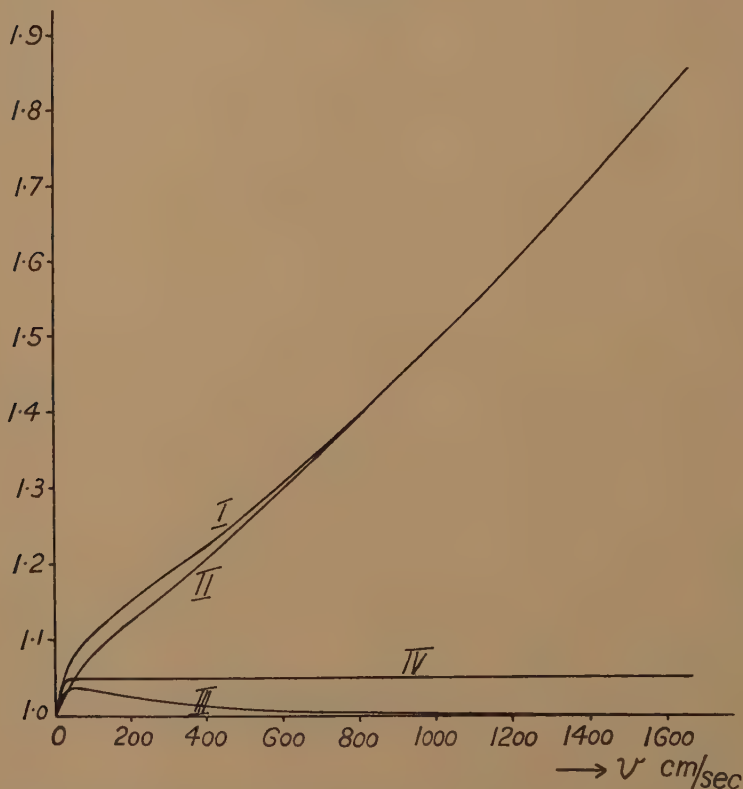
Corollaries.

(1) *The Variation of q with p .*—This is obtained directly from equations (13) and (22); to obtain identical distributions of diffusate in the stream of mercury vapour at different values of p ,

$$\frac{\pi^2 D}{4x_0^2} t_0 \propto \frac{T}{pv}$$

must be kept constant. Neglecting the slight difference in pressures between the gas mixture and the stream of mercury vapour, pv is proportional to the amount of diffusion medium Q_0 , passing the source of diffusion in unit time. As T increases with p , a bigger amount of diffusion medium, and hence a bigger rate of heating, is necessary to obtain the same distribution of diffusate and the same coefficient of fractionation

Fig. 7.



q is represented by curve I., q_l by II., q_h by III., and $\frac{\rho_s}{\rho_0}$ by IV.

at higher than at lower values of p . This effect is enhanced by the heat losses increasing with T . As shown in fig. 2 *b*, experiment and theory are in qualitative agreement.

As the increase of T with p is of less than first order in p to keep $\frac{T}{pv}$ constant, if p increases, v must decrease. This may be important in

practice, as the simplifying assumptions (a) and (c) made above are most likely to break down at low values of v (cf. further, paragraph (4) Cor. 2).

(2) *The Variation of q with D .*—This again is obtained directly from equation (13). If D_l/D_h has the same value for two pairs of gases, but the two D_l -values are different, to obtain identical distributions of diffusate ($D_l t_0$) must be given the same value in the two cases by adjusting t_0 . Hence the smaller D_l , the smaller the value of v , at which the distribution of diffusate in the stream of mercury vapour, corresponding to a given value of q , is obtained (cf. further, paragraph 4, Cor. 3.)

§4. Velocity of Fractionation.

Barwich's experiments for the determination of q do not give $q = \frac{\rho_{H\infty}}{\rho_{L\infty}}$

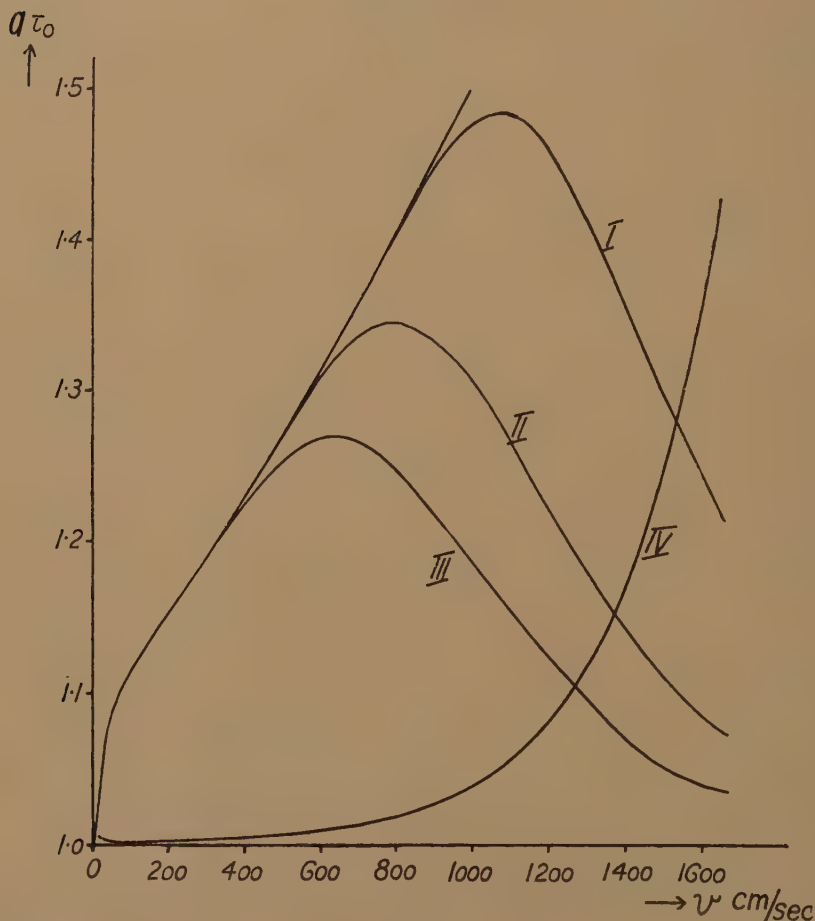
but $q_\tau = \frac{\rho_{H\tau}}{\rho_{L\tau}}$, where τ is a finite time constant throughout a series of experiments (cf. figs. (2 a) and (2 b)). To prove that $q_\tau \simeq q$ Barwich, in two experiments, increased τ sixfold (*loc. cit.*⁽³⁾, p. 177). They belong to series 1 (fig. 2 a); the rates of heating are ~ 190 and 260 watts respectively. At ~ 190 watts, no increase of q_τ is found, but at 260 watts q_τ increases from 3.5 to 5.2. Barwich concludes that the maximum value of $q = 14.5$ will not be reached at 260 watts, however big τ . This argument is faulty: firstly, we shall prove (cf. fig. 8, p. 358) that the velocity of separation decreases rapidly with v ; secondly, the strong decrease in the progress of separation from $\Delta q = 3.5$ in the first fifteen mins. to $\Delta q = 1.7$ in the next seventy-five mins. is probably only apparent. Barwich puts $t = 0$ when the light indicator sphere (cf. paragraph (3 b)) is lifted. By that time q has reached the value 1.5 already (*loc. cit.*⁽³⁾, fig. 9), though v is quite small. Before the mercury stream has reached its final velocity a few minutes must have passed, during which the rate of separation is much faster than at the final velocity. Hence probably only a small part of the $\Delta q = 3.5$ is achieved at the final value of v , which persists during the final seventy-five minutes in the prolonged experiment.

The velocity of fractionation of a Hertz apparatus consisting of several units in series has been calculated by Barwich and in a recent paper by Sherr. Barwich's calculation is carried out for the case of the concentration in one of the end volumes being constant, Sherr's for the general case of the amount of either component in the system being constant, and, hence, the concentration in both end volumes varying with time. In both calculations one component is assumed to be rare. Both solutions are complex. This complexity is due to more than one pump taking part in the fractionation. Hence it seems advantageous for a calculation of the change in the velocity of separation with v to develop a special

theory for the case of the single pump, using the picture of the process of separation described in paragraph 2 instead of the one used by Barwich.

As we are only concerned with the dependence of the velocity of separation on v , we can introduce the most convenient conditions as regards

Fig. 8.



Curves I., II., III., show the values of $q\tau_0$ for three values of τ_0 : $\tau_0 \propto 3.2$ for curve I., $\tau_0 \propto 1.0$ for II., $\tau_0 \propto .47$ for III. Curve IV. represents the variation in τ with v .

the concentrations in the apparatus. We choose $\rho_L = \text{constant}$ and $h < l$, which fits the experimental conditions for the enrichment of H_2O^{18} in water vapour⁽³⁾. For the same reason we can choose the case in which the equilibration between the heavy fraction and the gas in the heavy

side is complete (*cf.* paragraph 2), *i. e.*, the rate of equilibration is very big. This involves the assumption that the rate of equilibration only influences the velocity of fractionation, and not the dependence of the velocity of fractionation on v as well. This assumption is likely to be true, but we cannot prove it. Not so much, however, depends on its validity, because, as explained in paragraph 2, the equilibration does not only proceed by diffusion but also by irregular convection currents, and the rate of equilibration is probably fairly big.

Barwich's convention for fixing the zero of time in his experiments has been explained above; in the theoretical deduction below we assume that the separation starts at $t=0$, but that the stream of mercury vapour has already reached its final velocity at that time. The enrichment of the gas in the heavy side (of amount M_H gm.mols.) in the heavy component is due to a transport of heavy component from the source of constant isotopic concentration ρ_L in the light side to the source of diffusion, and from there into the heavy side. G_l gm.mols. per sec. of composition ρ_o/q_l flow from the source of diffusion to the source of constant isotopic composition in the light side, and G_l gm.mols./sec. of composition ρ_L return to the source of diffusion. Hence, using $h \ll l$, the net transport of heavy component from the light side to the source of diffusion is

$G_l \left(\rho_L - \frac{\rho_o}{q_l} \right)$ gm.mols./sec. G_h gm.mols. of gas of composition $\rho_o q_h$ flow from the source of diffusion to the heavy side, and G_h gm.mols./sec. of composition ρ_H return to the source of diffusion. Hence the net transport from the source of diffusion to the heavy side is $G_h(\rho_o q_h - \rho_H)$ gm.mols./sec. Neglecting the volume of the light side, we obtain

$$G_l \left(\rho_L - \frac{\rho_o}{q_l} \right) = G_h(\rho_o q_h - \rho_H) \quad \dots \dots \dots (25)$$

As, due to this transport, the amount of heavy component in the heavy side increases by $M_H \frac{d\rho_H}{dt}$ gm.mols./sec., we obtain

$$M_H \frac{d\rho_H}{dt} = G_l \left(\rho_L - \frac{\rho_o}{q_l} \right) \quad \dots \dots \dots (26)$$

and

$$M_H \frac{d\rho_H}{dt} = G_h(\rho_o q_h - \rho_H) \quad \dots \dots \dots (27)$$

Substitution of ρ_o from equation (26) in (27) leads to

$$\frac{d}{dt} \rho_H + \frac{G_h H_l}{M_H H_o} \rho_H = \frac{G_h H_l}{M_H H_o} q \rho_L \quad \dots \dots \dots (28)$$

using the assumption $l \gg h$ in the form $L \simeq G$ and equations (6), (7), (8), (22 a). In connexion with the boundary conditions $\rho_H = \rho_L$, if $t=0$, this gives

$$\rho_H = \rho_L \left\{ q - (q-1)e^{-\frac{G_h H_l}{M_H H_o} t} \right\} \quad . \quad . \quad . \quad . \quad . \quad (29)$$

Substituting equation (29) into (25) we obtain for :

$$\rho_o = \rho_L \left\{ q_l - (q_l-1)e^{-\frac{G_h H_l}{M_H H_o} t} \right\} \quad . \quad . \quad . \quad . \quad . \quad (30)$$

Let τ be the time in which—at a given value of v — ρ_H in a Hertz apparatus increases by a definite percentage, 95 per cent. say, of the possible increase $(q-1)\rho_L$. Then according to equation (29)

$$\tau \propto \frac{M_H H_o}{G_h H_l}.$$

As the amount of gas enriched during the time τ is M_H we define the velocity of fractionation by

$$\frac{M_H}{\tau} \propto \frac{G_h H_l}{H_o}.$$

The values of τ are shown in fig. 8, the scale along the y -axis being arbitrary.

Evaluation of equation (29) for a fixed value of t , τ_o , and for the values of $\frac{H_l}{G_h H_o}$ appropriate to the different values of v gives values of the density $\rho_{H\tau_o}$ obtained at different values of v , if the duration of fractionation is τ_o in every case. In fig. 8 $q_{\tau_o} = \frac{\rho_{H\tau_o}}{\rho_L}$ is plotted as function of v for three values of τ_o . Fig. 8 shows that in accordance with the experimental results theory predicts a decrease of q_{τ_o} if v increases beyond a certain value.

The experimental curves of figs. 2 a and 2 b show fairly equal gradients for the initial ascending and final descending parts. We can test whether the theoretical curves show similar characteristics, using $q=1.23$. This value has been obtained by interpolation by Barwich (*loc. cit.*, p. 190) and experimentally by J. B. M. Herbert⁽¹²⁾ in these laboratories. (Private communication.) The theoretical calculations are based on $D_l/D_h=1.1$, whereas the actual value of D_l/D_h for H_2O^{16} and H_2O^{18} is between 1.03 and 1.09. The variation of q with D_l/D_h ranges from $q \approx 1 + 10(D_l/D_h - 1)$ at $v \approx 15$ cm./sec. to $q \approx \exp. 5(D_l/D_h - 1)$ at $v \approx 1.400$ cm./sec. Hence the theoretical value of q corresponding to $q_{\text{exptl.}} = 1.23$ is between 1.3 and 1.7. Curve I (fig. 8) having its maximum at $q \approx 1.48$, has equally steep ascending

and descending branches, hence quantitative agreement between experiment and theory on this point is likely.

The theory thus explains the rapid decrease of q_{τ_0} for large values of v . The occurrence of a minimum of q under certain experimental conditions is discussed in the next paragraph.

Corollaries.

(1) *Velocity of Fractionation by a Series of, or a Single, Hertz Pump.*—It is of interest to compare the factors determining the velocity of separation in the case of a Hertz apparatus comprising only one unit, as given by equation (29), with those in the case of several units in series as calculated by Barwich. In the latter case (*cf.*, *loc. cit.*⁽³⁾, p. 119) the velocity of fractionation for a given number of pumps depends mainly on G_l (in our notation). According to equation (29), in the case of a single pump it is $\frac{G_h H_l}{H_o}$. As the difference between $\frac{H_l}{H_o}$ and $\frac{L_l}{L}$ is negligible compared with the rate of change of either of these ratios with v , we obtain :

$$\frac{G_h H_l}{H_o} \approx \frac{G_h L_l}{L_o} \approx \frac{L_h}{L_o} G_l \dots \dots \dots (31)$$

using $L \approx G$. Fig. 6 shows that $L_h/L_o \approx 1$ above small values of v and q , so that the velocities of separation in a Hertz apparatus comprising a single or several pumps show the same characteristic dependence on G_l .

(2) *Velocity of Fractionation for Different Values of p .*—As shown in paragraph 3 (c), if the pressure p in the apparatus is changed, and v is varied simultaneously, so that $\frac{T}{vp}$ is kept constant, the value of q remains unaltered. The effect on the velocity of fractionation is as follows : according to equations (20) and (21) L_o , L_v , etc., are proportional to vl_s if $\frac{T}{vp}$ is constant. As $l_s \propto p$,

$$vl_s \propto vp \propto T.$$

Hence the velocity of fractionation $\frac{G_h H_l}{H_o}$, measured in gm.mols./sec., is proportional to T , *i. e.*, increases slightly with p . As, however, for a given apparatus $M_H \propto p$, $\tau \propto \frac{p}{T}$. As the increase of T with p is of less than first order, τ increases, though a greater mass of gas is fractionated per sec.

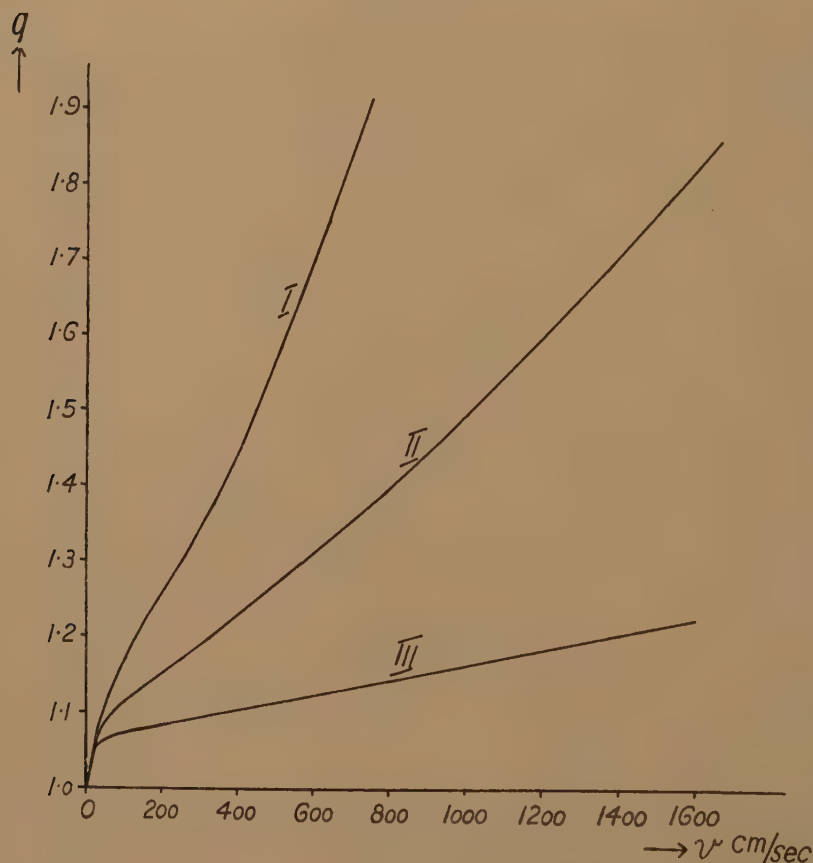
(3) *Velocity of Fractionation for Different Values of D .*—As shown in paragraph 3 (c), the value of v , at which a given value of q is obtained

decreases, if D increases. Hence L_0 , L_v , etc., decrease as well, and with them the velocity of fractionation.

§ 5. The Minimum of $q_{\text{exptl.}}$

In paragraphs (3) and (4) we have assumed that $x_h = \frac{1}{3}x_0 = .3$ cm. for all values of v (paragraph 3 (a), assumption (b)). To test the importance

Fig. 9.



Variation of q with v at various values of $\frac{x_h}{x_0}$, if $x_0 = 0.9$ cm. Curve I: $\frac{x_h}{x_0} = \frac{1}{2}$; curve II., $\frac{1}{3}$; curve III., $\frac{1}{6}$.

of this assumption, q and τ are calculated for two more values of x_h , $x_h = .45$ cm. $= \frac{1}{2}x_0$ and $x_h = .15$ cm. $= \frac{1}{6}x_0$, substituting the appropriate equations

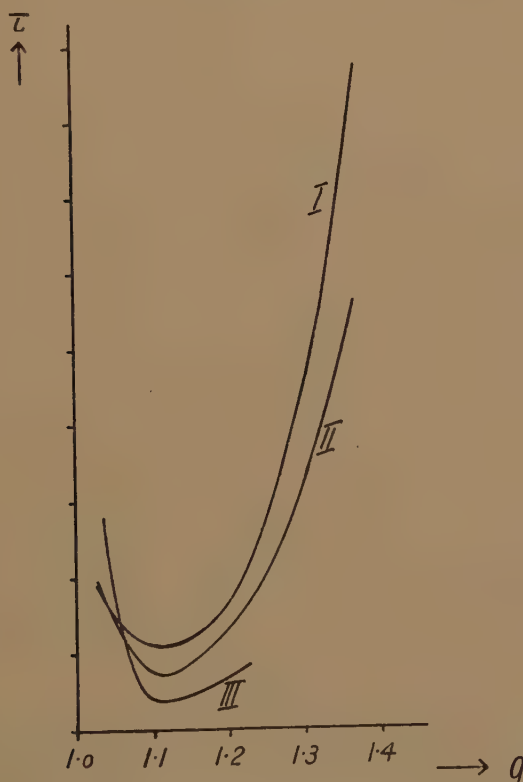
$$l_l = l_s \left\{ 1 - \frac{48}{5\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \cos \frac{(2n+1)\pi}{12} e^{-(2n+1)^2 \pi l_0} \right\}. \quad (32)$$

in the case of $x_h = 0.15 \text{ cm.} = \frac{1}{3}x_0$, and

$$l_l = l_s \left\{ 1 - \frac{16}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \cos \frac{(2n+1)\pi}{4} e^{-(2n+1)^2 f_0} \right\} \quad (33)$$

in the case of $x_h = 0.45 \text{ cm.} = \frac{1}{2}x_0$,—for equation (19). The results are plotted in figs. 9 and 10. They show that above very small values of v , q is exceedingly sensitive to variations in the value of x_h .

Fig. 10.



τ plotted as function of q for different values of $\frac{x_h}{x_0}$, $x_0 = 0.9 \text{ cm.}$ Curve I.:

$\frac{x_h}{x_0} = \frac{1}{2}$; curve II. : $\frac{1}{3}$; curve III. : $\frac{1}{6}$. The scale along the y -axis is arbitrary.

The only safeguard for the constancy of x_h in the Hertz pump is the edge PP. Its approximate shape, as it mostly occurs in the diffusion pumps, is shown in fig. 1 *b*. Quite apparently this is not sufficiently definite to keep x_h independent of v , if the pressure of the mercury vapour

at A varies with v without being counteracted equally at P_1 and F_1 by equal changes in pressure.

The pressure at A must be sufficient to drive the mercury vapour and the light fraction up the tube P_1P_2 against gravitational and viscous resistances. The gravitational resistance, R_g , is mainly due to the diffusion medium, the viscous resistance, R_v , mainly due to the diffusate. R_g is roughly proportional to the amount of mercury vapour going up the tube P_1P_2 , and R_v to the amount of gas in the light fraction. Hence in rough approximation

$$R_g \propto v, \quad (34)$$

$$R_v \propto G_l \approx L_l. \quad (35)$$

Hence R_g increases linearly with v , whereas R_v attains a maximum at a fairly low value of v , v_L , and then decreases. If a capillary is introduced into the path of the light fraction, R_v increases. This increase is made up of two parts: the resistance of the capillary against the flow of the light fraction through it, and the resistance of the excess pressure of diffusate, built up in front of the capillary, against the flow of mercury up the tube P_1P_2 .

The relative magnitudes of R_v and R_g cannot be estimated; we shall assume—and later justify by comparing with experiment the inferences drawn from this assumption—that $R_v > R_g$ for values of v in the neighbourhood of v_L , even if no capillary resistance is inserted into the tube P_1P_2 .

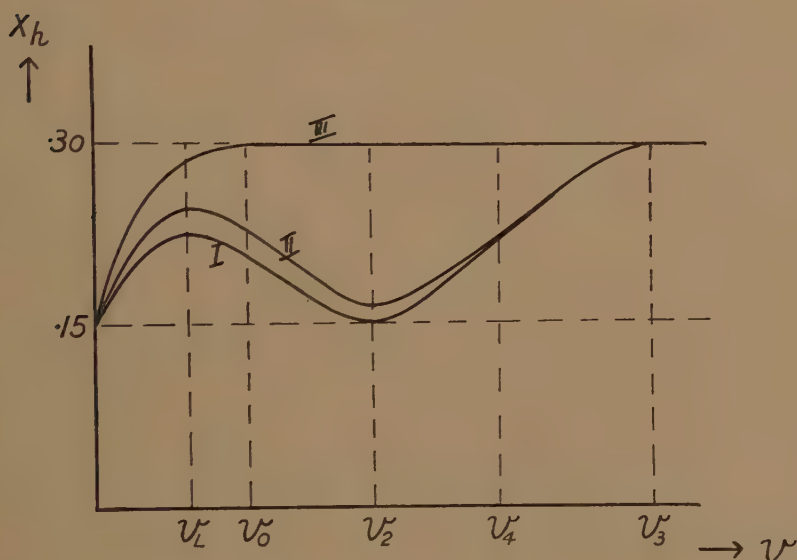
Hence the variations of $(R_v + R_g)$ with v are: from $v=0$ to $v=v_L$ $(R_v + R_g)$ increases. On increasing v beyond v_L , R_v decreases quicker than R_g increases because G_l decreases quicker than v increases (cf. fig. 6). Hence $(R_v + R_g)$ decreases above v_L . This decrease stops at v_2 , say, when R_v becomes too small for the increase of R_g to make up for the decrease of R_v . Hence a minimum of $(R_v + R_g)$ is obtained at v_2 , followed by a roughly linear increase.

The energy necessary to bring the outer part of the mercury stream and the heavy fraction to F_1 is small for all values of v : the flow is nearly horizontal, condensation of the diffusion medium quicker than in P_1P_2 due to the bigger ratio between condensing surface and condensate, and the viscous resistance small due to the width of the annulus F_1F_2 .

So the changes of pressure at A due to the variations of $(R_v + R_g)$ with v are not counterbalanced by equal variations of the resistance at F_1 . Hence if $(R_v + R_g)$ is large, a part of the mercury stream and diffusate will be pressed towards F_1F_2 , which, if the position of the edge PP were solely decisive, would go up P_1P_2 . Hence x_h varies with v . These variations will only take place inside certain limits of x_h between $x_h = .3 \text{ cm.} = \frac{1}{3}x_0$ and $x_h = .15 \text{ cm.} = \frac{1}{8}x_0$, say.

Three sets of variations of x_h with v corresponding to the three sets of experimental conditions used by Barwick are sketched roughly in fig. 11. Barwick's fourth set of conditions (no capillary, but no connexion from V_L to the annulus C_1C_2) is not discussed here, as irregularities in the stream of mercury vapour allowing gas to pass from the light side to the source of diffusion are essential in this case (*cf.*, *loc. cit.*⁽²⁾). Curve 1 holds if no capillary is inserted into tube P_1P_2 . This leads to a flat maximum of R_o at v_L and to a low minimum of $(R_o + R_g)$ at v_2 , so that x_h attains its lower limiting value at v_2 . Curve 2 obtains if a capillary of low resistance (3 mm. wide, 5 cm. long) is inserted into the tube P_1P_2 . Curve 3, if the

Fig. 11.



capillary resistance in tube P_1P_2 is large (capillary 3 cm. long, 1 mm. wide). In that case $(R_o + R_g)$ is large enough to make x_h attain its higher limiting value even for $v = v_2$.

The variations of the coefficient of separation with v , under the conditions illustrated in fig. 11, are as follows:—Curve 3: at values of v greater than v_0 (*cf.* fig. 11) values of q calculated above (*cf.* fig. 7) for $x_h = 0.3$ cm. $= \frac{1}{3}x_0$ hold. The increase of x_h in the range $0 < v < v_0$ produced a slower ascent of q with v than in the case $x_h = 0.3$ cm. for the first part of this range, and causes it to be much steeper in the second part. This effect will be demonstrated in experimental results obtained by Sherr (*cf.* paragraph 6). Curves 1 and 2: range $0 < v < v_L$ similar to range $0 < v < v_0$ in the case of curve 3. In the range $v_L < v < v_2$ the value of x_h

decreases so quickly that the corresponding decrease of q exceeds the increase of q brought about by the increase in v . Fig. 10 shows that v_2 can be several times as big as v_L without the increase of q due to increase in v compensating the decrease of q due to the pronounced decrease of x_h . But being the result of two opposing influences the decline of q in the range $v_L < v < v_2$ is slow. If v increases beyond v_2 , q increases with v and x_h ; hence the rise of q in the range $v_2 < v < v_3$ is more rapid than in the case $x_h = .3 \text{ cm.} = \frac{1}{3}x_0$. Hence the coefficient of separation under the condition of curves 1 and 2 has a maximum at a fairly low value of v , followed by a minimum at higher values of v . The values of the maxima and minima of curves 1 and 2 in fig. 11, on comparison with the results of fig. 9, show that in the range $0 < v < v_4$ (cf. fig. 11) higher values of q are obtained under the conditions of curve 2, fig. 11, than under those of curve 1. The maxima and minima of q so obtained correspond to the first maxima and the minima of the experimental curves of fig. 2 *a*, curves 1 and 2. Thus the theory not only explains the existence of these stationary values, but also gives the reasons why the minimum and maximum are obtained at higher values of q under the conditions of curve 2 than under those of curve 1, fig. 2 *a*, and why the minima of both curves are unsymmetrical in shape, being flat towards lower values of v and very steep towards higher values.

Hence, the assumption that $R_v > R_q$ for values of $v \approx v_L$, leads to the inference that under the conditions of curves 1 or 2, fig. 2 *a*, a maximum of q is obtained at v_L . But according to fig. 2 *b*, this first maximum is obtained at lower values of v than the maximum of Δp . In paragraph 3 (*b*), Cor. 2, we explained that Δp is roughly proportional to $L_t L_h$, and fig. 6 shows that the maximum of L_t occurs at a lower value of v than that of $L_t L_h$. Hence the fact that the first maximum of q is obtained at a lower value of v than the maximum of Δp is a strong support for the conjecture that the maximum of q is produced by the maximum of L_t , i. e., G_t . This is only possible if the assumption $R_v > R_q$ at values of $v \approx v_1$ is correct.

To complete the theoretical reproduction of the results of fig. 2 *a* for $q_{\text{exptl.}}$, we take account of the fact that the fractionation is obtained in a finite time. According to fig. 8, the velocity of fractionation decreases with v for any value of x_h . Hence the theory gives a maximum followed by a decrease of $q_{\text{exptl.}}$ at high values of v . The experimental results show that under the condition of curve 3 the maximum of q is attained at a velocity not much bigger than v_L . For the value of v , at which the second maxima of curves 1, 2, fig. 2 *a*, are obtained, curve 3 gives a small value of q . Hence these maxima are obtained at values of x_h smaller than .3 cm. Fig. 10 shows that τ decreases simultaneously with x_h for a fixed value of q . This is the reason why nearly the same value of q can

be obtained in twenty minutes under the experimental conditions of curves 1 or 2, fig. 2 *a*, as in forty minutes under those of curve 3.

Hence we have been able, by introducing into the theoretical considerations perfectly regular, non-turbulent aerodynamic effects which lead to variations in the magnitude of x_h , to explain the occurrence of a minimum in the curves for q_{exptl} .

§6. Improvement of Hertz's Diffuser.

In a recent paper, Sherr describes attempts to construct a modification of the Hertz diffuser, more efficient than the original. Before discussing his methods and results we give a summary of the possibilities of improving Hertz's apparatus suggested by our theory.

(1) If x_h , the radial distance between the source of diffusion and the effective separating edge, is made independent of v , the minimum of q disappears, the necessity for inserting a narrow capillary into the tube P_1P_2 , which cuts down the flow of diffusate from pump to pump, is obviated and the maximum of q_{τ_0} , even compared with the one obtained at the cost of inserting the narrow capillary, broadened (*cf.* paragraph 5, and discussion of Sherr's results below).

(2) As shown in fig. 10, τ decreases simultaneously with x_h for a fixed value of q . Hence the smaller x_h , the greater the speed of fractionation.

(3) Comparing fig. 8 with fig. 9 we see that the slow rise of q with v at low values of x_h means a broader maximum of q_{τ_0} at lower than at higher values of x_h .

(4) As proved at the end of paragraph 4 the mass of gas fractionated in unit time slightly increases with p for a given value of q (*cf.* paragraph 3 (*b*)).

(5) In modifying Hertz's model of the diffuser, care must be taken not to disturb the mixing between the heavy fraction produced in one unit and the light fraction of the next "heavier" unit as described in paragraph 2.

Sherr's Experiments.

Observing the stream of mercury vapour in a Hertz pump by means of a high frequency discharge, Sherr found that the jet formed above BB is very ill-defined. To obtain a jet of better definition the tube P_1P_2 is connected to the tube FF by means of a ring-seal and continued below this ring-seal, so that the end of a tube of about 1 mm. wall-thickness serves as edge PP. The radial clearance between the tubes P_1 and B_1 is made 1 mm. in one set of pumps (type B) and 2 mm. in another set (type C). 12 Hertz pumps, 9 pumps of type B, and 8 pumps of type C are connected in series and all the test experiments carried out on this

composite apparatus, observing the separation effected by each set of pumps separately. The coefficients of separation for a set of pumps are reduced to the coefficients for single pumps by taking the R th root, R being the number of pumps in the set. A narrow capillary is inserted into the tube P_1P_2 of every pump. Unfortunately, Sherr does not state the duration of the experiments, but it can be assumed that all the runs were of similar lengths. Sherr's results confirm our theory in the following points :—

(a) In spite of the better defined jet of mercury-vapour $q_{\text{exptl.}} = q_{\tau_0}$ decreases if v is increased beyond a certain maximum value.

(b) The maximum value of q_{τ_0} is obtained at a far higher value of the rate of heating, *i. e.*, v , in type B, than in type C. As in the case of the Hertz pump the maximum is obtained at a still higher value of v , we infer that for these pumps $x_h < 1$ mm.

(c) The maximum of q_{τ_0} is broader for type B than type C.

(d) The maximum of q for the Hertz pump is narrower than that of type B, due to the steeper increase of q with v (*cf.* §5). This difference is seen to be very pronounced.

Conclusions.

It seems probable that in Sherr's modification of the Hertz pump x_h is sufficiently independent of v to avert the occurrence of a minimum of q , if the pump is worked without a capillary constriction in the tube P_1P_2 . If worked without a capillary, pump B should show a speed of separation greater than that of pump C, as the above theory leads to expect. That this is not evident from the present experiments on the fractionation of neon is probably due to the influence of the capillary constriction (*cf.* paragraph 3, assumption (c)). For a final comparison it would also seem preferable to work the pumps B and C independently and not in series as in the present arrangement.

SUMMARY.

A simplified theory, suitable for numerical work, of the coefficient and the velocity of fractionation of Hertz's apparatus for the fractionation of isotopic gases by diffusion into a stream of mercury vapour is developed, assuming streamline motion of the vapour. In agreement with experiment it is shown that the decrease at high velocities of the vapour stream, of the fractionation obtained in a finite time, as observed by Barwich, can be explained by this theory, and hence is not due to any irregular phenomena as assumed before. The reason for a minimum of fractionation, observed under certain experimental conditions at moderate

velocities of the vapour stream, is found in perfectly regular aerodynamic effects, and means of overcoming these effects are given. The influence of the geometrical dimensions of the apparatus, and the pressure and the diffusion coefficient of the gas in the apparatus on the coefficient and rate of fractionation are discussed in detail, and the results compared with Barwich's and Sherr's experimental work.

Acknowledgment.

The author wishes to express his sincere thanks to Professor M. Polanyi and Mr. J. B. M. Herbert for much helpful comment, to Mr. O. Bünemann for a thorough criticism of the manuscript, and to his father, Professor O. Blumenthal, for advice in the mathematical part of the work.

References.

- (1) G. Hertz, *Z. f. Physik.* xix. p. 35 (1923); and Proc. Amsterdam, xxv. p. 434 (1923).
- (2) G. Hertz, *Z. f. Physik.* xci. p. 810 (1934).
- (3) H. Barwich, *Z. f. Physik.* c. p. 166 (1936).
- (4) D. MacGillavry, Trans. Far. Soc. xxxiii. p. 439 (1937).
- (5) R. Sherr, J. Chem. Phys. vi. p. 251 (1938).
- (6) Cf. M. Trautz, *Ann. d. Phys.* (5) xxii. p. 364 (1935).
- (7) H. v. Hartel, N. Meer, M. Polanyi, *Z. Phys. Chem.* B, xix. p. 145 (1932).
- (8) D. MacGillavry, "The Concentration of Hydrogen of Atomic Weight Two by Diffusion in Fast Streaming Vapours." Dissertation. Columbia University, New York City (1933).
- (9) E. L. Lederer, *Kolloid. Z.* xlv. p. 108 (1928).
- (10) Landolt-Börnstein. 5. Aufl. Hw. 250.
- (11) J. H. Jeans, 'The Dynamical Theory of Gases,' 4th Edition, p. 317.
- (12) Cf. E. Blumenthal and J. B. M. Herbert, Trans. Far. Soc. xxxviii. p. 849 (1937).

XXXII. *Entropy of Fermi-Dirac Gas.*

By F. C. AULUCK, Panjab University, Lahore *.

[Received August 4, 1938.]

ENTROPY of Fermi-Dirac gas is found up to the second order of approximation.

In recent years various applications of F.D. statistics have been made to physical and astrophysical problems. It is now an established fact that the theory of the degenerate electron gas (degenerate in the sense of Fermi-Dirac statistics) can satisfactorily explain the internal constitution of the white dwarf stars and planets ⁽¹⁾. The general properties of the degenerate and the non-degenerate gas have, therefore, been studied in detail during recent years by Fowler, Milne, Majumdar, Kothari, and others. In this paper I take one such property, that is, the entropy of F.D. electron gas. So far the expressions for entropy were not given up to the second order ⁽²⁾. This is done in this paper.

1. Consider an assembly of N free-electrons occupying a volume V . The number of wave functions for which the momentum s lies between s and $s+ds$ is known to be

$$V \frac{4\pi s^2 ds}{h^3} \dots \dots \dots (1)$$

But in the case of an electron

$$s = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}}, \dots \dots \dots (2)$$

and the kinetic energy

$$\epsilon = mc^2 \left\{ \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right\}, \dots \dots \dots (3)$$

where m is the rest mass of the electron, c is the velocity of light, h Planck's constant. Eliminating v from (2) and (3) we get

$$s^2 = \frac{1}{c^2} (\epsilon^2 + 2mc^2\epsilon). \dots \dots \dots (4)$$

* Communicated by Dr. D. S. Kothari.

Hence the number of wave functions for which the energy ϵ lies between ϵ and $\epsilon + d\epsilon$ is

$$a(\epsilon)d\epsilon = V \frac{4\pi g}{c^3 h^3} (\epsilon^2 + 2\epsilon mc^2)^{1/2} (\epsilon + mc^2) d\epsilon,$$

where g is the weight factor which is equal to 2 for an electron. Then the Fermi-Dirac distribution law states that the number $n(\epsilon)d\epsilon$ of electrons possessing kinetic energy between ϵ and $\epsilon + d\epsilon$ is given by

$$n(\epsilon)d\epsilon = \frac{a(\epsilon) d\epsilon}{\frac{1}{A} e^{\epsilon/kT} + 1}, \quad \dots \quad (5)$$

where A is independent of ϵ , k is Boltzmann's constant.

Hence

$$\begin{aligned} N &= \int_0^\infty n(\epsilon) d\epsilon \\ &= V \frac{4\pi g (kT)^3}{(ch)^3} \int_0^\infty \frac{(u^2 + 2uy)^{1/2} (u + y) du}{\frac{1}{A} e^u + 1}, \quad \dots \quad (6) \end{aligned}$$

where

$$u = \frac{\epsilon}{kT}, \quad y = \frac{mc^2}{kT}.$$

From the general principles of statistical mechanics it is known that the entropy S is given by

$$S = -kN \log A + \frac{1}{3} kV \frac{4\pi g}{c^3 h^3} (kT)^3 \int_0^\infty \frac{d\{u(u^2 + 2uy)^{3/2}\}}{\frac{1}{A} e^u + 1}. \quad \dots \quad (7)$$

It is difficult to evaluate the integrals (6) and (7) for an arbitrary value of A , but we shall do it in two important limiting cases—Case (i.)—the non-degenerate case, where $A \ll 1$. In this case the behaviour of the gas approximates to that of the *classical* perfect gas. Case (ii.)—the degenerate case, where $A \gg 1$. In this case the behaviour of the gas departs widely from that of the *classical* perfect gas. These are further subdivided into two cases—relativistic and non-relativistic. In the former case the kinetic energy per electron is very large compared to its rest mass energy, and for the other it is just the opposite.

2. Non-Degenerate Case $A \ll 1$.

(i.) *Non-relativistic*.—In this case $\frac{kT}{mc^2} \ll 1$.

Therefore (6) reduces to

$$n = \frac{N}{V} = 4\pi g \left(\frac{kT}{ch} \right)^3 2^{1/2} y^{3/2} \int_0^\infty \frac{u^{1/2} + \frac{5}{4} y^{-1} u^{3/2}}{\frac{1}{A} e^u + 1} du$$

$$= 4\pi g \left(\frac{kT}{ch} \right)^3 2^{1/2} y^{3/2} A \left\{ \frac{\sqrt{\pi}}{2} \left(1 - \frac{A}{2^{3/2}} + \dots \right) + \frac{15}{16y} \sqrt{\pi} \dots \right\}^*$$

Hence

$$A = A_0 \left\{ 1 - \frac{15}{8y} + \frac{A_0}{2^{3/2}} \right\}, \quad \dots \dots \dots (8)$$

where

$$A_0 = \frac{n\hbar^3}{g(2\pi mkT)^{3/2}}.$$

Again the expression for entropy reduces to

$$S = -kN \log A + \frac{1}{3} \frac{kV \cdot 4\pi g}{(ch)^3} (kT)^3 (2y)^{3/2} \int_0^\infty \frac{\frac{5}{2} u^{3/2} + \frac{21}{8y} u^{5/2}}{\frac{1}{A} e^u + 1} du$$

$$= -kN \log A_0 + kN \left\{ \frac{5}{2} + \frac{15}{4y} + \frac{A_0}{2^{7/2}} \right\}, \quad \dots \dots \dots (9)$$

omitting powers of y higher than the first.

(ii.) *Relativistic case.*—In this case

$$\frac{kT}{mc^2} \gg 1.$$

Hence, keeping only the first power of y , we get

$$n = 4\pi g \left(\frac{kT}{ch} \right)^3 \int_0^\infty \frac{u^2 + 2uy}{\frac{1}{A} e^u + 1} du.$$

Integrating it and solving for A we get

$$A = A_0^R \left(1 - y + \frac{A_0^R}{8} \right), \quad \dots \dots \dots (10)$$

where

$$A_0^R = \frac{n}{8\pi g} \left(\frac{ch}{kT} \right)^3.$$

* We get it by using the result

$$\int_0^\infty u^n e^{-ru} du = \frac{1}{r^{n+1}} \Gamma(n+1).$$

Similarly, it can be shown that to the second degree of approximation

$$S = kN \left\{ -\log A_0^R + 4 + \frac{A_0^R}{8} \right\} \dots \dots \dots (11)$$

3. Degenerate Case $A \gg 1$.

We apply Sommerfeld's⁽³⁾ result : if $\phi(u)$ is a regular function then

$$\int_0^\infty \frac{d\phi(u)}{\frac{1}{A}e^u + 1} = \phi(u_0) + \frac{\pi^2}{6} \phi''(u_0) + \frac{7\pi^4}{360} \phi^{iv}(u_0) + \dots,$$

where $u_0 = \log A$.

(i.) We first deal with the general case when

$$x = \frac{h}{m^2} \left(\frac{3n}{4\pi g} \right)^{1/3}$$

has any value.

Now

$$n = \frac{4\pi g}{3} \left(\frac{kT}{ch} \right)^3 \int_0^\infty \frac{d(u^2 + 2uy)^{3/2}}{\frac{1}{A}e^u + 1},$$

or

$$(xy)^3 = \int_0^\infty \frac{d(u^2 + 2uy)^{3/2}}{\frac{1}{A}e^u + 1} \dots \dots \dots (12)$$

Putting $z^2 = u_0^2 + 2u_0y$ and applying Sommerfeld's method, we get

$$(xy)^3 = z^3 + \pi^2 \left(z + \frac{y^2}{2z} \right) + \frac{7\pi^4 y^4}{40 z^5} + \dots, \dots \dots (13a)$$

which gives

$$z = xy - \frac{\pi^2}{3xy} - \frac{\pi^2 y^2}{6(xy)^3} + \left(\frac{\pi^6}{81} - \frac{\pi^4 y^2}{9} \right) \frac{1}{(xy)^5} \dots \dots (13b)$$

Since

$$z^2 = u_0^2 + 2u_0y$$

therefore

$$\log A = \sqrt{y^2 + z^2} - y. \dots \dots \dots (14)$$

For finding the entropy we integrate (7) by the above method, and using the result (13 a) we get

$$S = \frac{1}{3} kV 4\pi^3 g \left(\frac{kT}{ch} \right)^3 \sqrt{y^2 + z^2} \left\{ z + \frac{7\pi^2}{15z} - \frac{7\pi^2 y^2}{30z^3} \right\}, \dots \dots (15)$$

correct to the second order of approximation

$$= \frac{R\pi^2}{(xy)^3} \sqrt{y^2 + z^2} \left\{ z + \frac{7\pi^2}{15z} - \frac{7\pi^2 y^2}{30z^3} \right\},$$

where

$$R = kN.$$

To get the first approximation we take $z=xy$, then

$$\log A = \frac{mc^2}{kT} \{ \sqrt{1+x^2} - 1 \}$$

and

$$\begin{aligned} S &= \frac{R\pi^2}{x^3y} x \sqrt{1+x^2} \\ &= \frac{4g\pi^3 k^2 V m^2 c}{3h^3} \{ x(1+x^2)^{1/2} T \}. \end{aligned}$$

(ii.) The two subcases, completely relativistic and completely non-relativistic, can be easily obtained from (14) and (15) with the help of (13 b).

For the completely non-relativistic case $x \rightarrow 0$, and therefore

$$\log A = \frac{yx^2}{2} \left\{ 1 - \frac{\pi^2}{3y^2x^4} - \frac{1}{4}x^2 \right\}, \quad \dots \quad (16)$$

$$S = \frac{R\pi^2}{yx^2} \left\{ 1 - \frac{2\pi^2}{5y^2x^4} + \frac{1}{2}x^2 \right\}, \quad \dots \quad (17)$$

where

$$\frac{yx^2}{2} = \frac{h^2}{2mkT} \left(\frac{3n}{4\pi g} \right)^{2/3} = \left(\frac{3A_0 \sqrt{\pi}}{4} \right)^{2/3}.$$

For the completely relativistic case $x \rightarrow \infty$, and therefore

$$\log A = xy \left\{ 1 - \frac{\pi^2}{3x^2y^2} - \frac{1}{x} \right\}, \quad \dots \quad (18)$$

$$S = \frac{R\pi^2}{xy} \left\{ 1 - \frac{\pi^2}{5x^2y^2} + \frac{1}{2x^2} \right\}, \quad \dots \quad (19)$$

where

$$xy = \frac{hc}{kT} \left(\frac{3n}{4\pi g} \right)^{1/3} = (6A_0^R)^{1/3}.$$

References.

- (1) Kothari, Mon. Not. R. Astro. Soc. xevi. p. 833 (1936).
- (2) Majumdar, *Astr. Nachr. Mr.* 5916-17, Band 247, p. 250 (1932); D. S. Kothari and Singh, *Zeit. für Astr.* xv. p. 146 (1938).
- (3) Sommerfeld, *Zeit. für Physik*, xlvii. pp. 1-32 (1928).

XXXIII. *Hamilton's Quaternions and Minkowski's Potentials.*

By OTTO F. FISCHER *.

[Received November 1, 1938.]

CONTENTS.	Page
1. Introduction	375
2. Hamilton's Quaternions	376
3. Maxwell's Equations	376
4. Minkowski's System	377
5. Klein's Hint	378
6. My Plan	378
7. The Four-dimensional Differentiation	379
8. The Primary Potential ψ	380
9. The Conditions of the First Step: the "A-system" ..	381
10. Minkowski's Potentials	382
11. Have Minkowski's Assumptions been too hasty? ..	383
12. Conclusion	385

1. *Introduction.*

ENGLISH scientific literature has produced an extraordinarily beautiful mathematical tool in Hamilton's Quaternions. In no other country of course this is so well known as in England itself. If it appears out of place then that I—a Scandinavian—should write of Hamilton's Quaternions in an English article let it be put down as a tribute to the great interest of his creation even to other countries.

In spite of their mathematical perfection, however, Hamilton's Quaternions have been very little used outside abstract mathematics. Is it not unsatisfactory to possess a fine tool, on whose perfection very much ingenious work has been laid down, and not use it?

Why, for instance, have the four-dimensional quaternions not been taken up as a means of formulating and treating modern four-dimensional universe?

The answer as a rule is: we have tried, but not succeeded. Minkowski himself goes so far as to declare this road impassable.

Nevertheless a few signs remain that the road is less impassable than indicated. In this article I hope to be able to reinforce such hints and take my first step towards a suggestion of more general use of this beautiful tool, which inspires new points of view.

* Communicated by the Author.

After a few words on terms and nomenclature I shall take my starting-point from Maxwell's equations.

2. *Hamilton's Quaternions.*

Altogether passing over the philosophy of Hamilton's Quaternions, the closed group of their system and their origin as vector fractions, I confine myself very shortly to the laying down of a few simple terms for use in this article. I shall define quite primitively a quaternion as the sum of a scalar and an ordinary three-dimensional vector, whereby the rule of scalar multiplication of two vectors is different so far that the sign is altered (like the negative result of squaring the imaginary unit).

Using italics for the synthetic term I shall write a complete quaternion as follows :—

$$O = \mathbf{O} + \sigma = \mathbf{O}_1 \mathbf{i}_1 + \mathbf{O}_2 \mathbf{i}_2 + \mathbf{O}_3 \mathbf{i}_3 + \sigma, \quad (1)$$

in the last expression specifying the vector coordinates corresponding to the unit vector axes $\mathbf{i}_1 \mathbf{i}_2 \mathbf{i}_3$.

A complete quaternion multiplication, corresponding to the sum of a scalar multiplication and a vector multiplication, I shall indicate and specify as follows :—

$$\begin{aligned} O \times F &= O \times F + O . F \\ &= \mathbf{O} \times \mathbf{F} - \mathbf{O} . \mathbf{F} + \sigma . f + \sigma . \mathbf{F} + \mathbf{O} . f \\ &= \mathbf{i}_1 (\sigma \mathbf{F}_1 + f \mathbf{O}_1 + \mathbf{O}_2 \mathbf{F}_3 - \mathbf{O}_3 \mathbf{F}_2) \\ &\quad + \mathbf{i}_2 (\sigma \mathbf{F}_2 + f \mathbf{O}_2 + \mathbf{O}_3 \mathbf{F}_1 - \mathbf{O}_1 \mathbf{F}_3) \\ &\quad + \mathbf{i}_3 (\sigma \mathbf{F}_3 + f \mathbf{O}_3 + \mathbf{O}_1 \mathbf{F}_2 - \mathbf{O}_2 \mathbf{F}_1) \\ &\quad + \sigma f - \mathbf{O}_1 \mathbf{F}_1 - \mathbf{O}_2 \mathbf{F}_2 - \mathbf{O}_3 \mathbf{F}_3. \quad (2) \end{aligned}$$

3. *Maxwell's Equations.*

Maxwell's fundamental electromagnetic equations I shall write in the following way :—

A : *for empty space* (the "A-system" including electric current) :

$$-\text{curl } \mathbf{E} = \frac{1}{c} \dot{\mathbf{H}}, \quad (\text{A I})$$

$$\text{curl } \mathbf{H} = \frac{1}{c} (\rho \mathbf{v} + \dot{\mathbf{E}}), \quad (\text{A II})$$

$$\text{div } \mathbf{H} = 0, \quad (\text{A III})$$

$$\text{div } \mathbf{E} = \rho, \quad (\text{A IV})$$

where \mathbf{E} denotes the electromotive force, \mathbf{H} the magnetic force, c the velocity of light, ρ the electric specific charge, and \mathbf{v} its velocity, constituting an electric current $\mathbf{J} = \rho \mathbf{v}$.

B : for matter (the "B-system") :

$$-\text{curl } \mathbf{E} = \frac{1}{c} \dot{\mathbf{B}}, \quad \dots \dots \dots \quad (\text{BI})$$

$$\text{curl } \mathbf{H} = \frac{1}{c} (\mathbf{J} + \dot{\mathbf{D}}), \quad \dots \dots \dots \quad (\text{BII})$$

$$\text{div } \mathbf{B} = 0, \quad \dots \dots \dots \quad (\text{BIII})$$

$$\text{div } \mathbf{D} = \rho, \quad \dots \dots \dots \quad (\text{BIV})$$

where **D** denotes the dielectric displacement and **B** the magnetic induction. Differentiation with regard to the time *t* is indicated by a point above the letter in question.

4. Minkowski's System.

The interpretation of Maxwell's equations by Minkowski's system of four-dimensional and six-dimensional vectors and potentials, which has so fundamentally influenced modern physics, may shortly be described as follows :—

a. A "fourth dimension" is introduced, characterized by the co-ordinate :

$$l = ict = i\mu \quad \mu = ct, \quad \dots \dots \dots \quad (3)$$

where *i* denotes the imaginary unit.

b. A four-dimensional radius vector is introduced, composed of the ordinary radius vector,

$$\mathbf{r} = x\mathbf{i}_1 + y\mathbf{i}_2 + z\mathbf{i}_3, \quad \dots \dots \dots \quad (4)$$

and the fourth dimension (3).

c. A four-dimensional potential Ω is introduced, composed in corresponding manner of a vector potential **A** and a scalar potential *iφ* defined by

$$\mathbf{H} = \text{curl } \mathbf{A}, \quad \dots \dots \dots \quad (5)$$

$$\text{curl } \left(\frac{1}{c} \dot{\mathbf{A}} + \mathbf{E} \right) = 0, \quad \dots \dots \dots \quad (6)$$

$$-\mathbf{E} = \frac{1}{c} \dot{\mathbf{A}} + \nabla \phi, \quad \dots \dots \dots \quad (7)$$

and the assumption of the following dependence :

$$\frac{1}{c} \dot{\phi} + \nabla \mathbf{A} = 0. \quad \dots \dots \dots \quad (8)$$

d. A four-dimensional current **P** is introduced, composed of the vector

$\frac{\rho}{c} \mathbf{v}$ and the scalar ρ .

e. A new four-dimensional divergence operator is introduced along the line

$$\text{Div } \Omega = \sum_{1 \text{ (k)}}^4 \frac{\delta \Omega_k}{\delta x_k} \dots \dots \dots (9)$$

f. A new "six-dimensional curl operator" along the line

$$\text{Curl}_{kh} \Omega = \frac{\delta \Omega_h}{\delta x_k} - \frac{\delta \Omega_k}{\delta x_h} : k, h = 1, 2, 3, 4. \dots \dots (10)$$

g. A new "four-dimensional Laplace-operator" is introduced :

$$\square = \sum_{1 \text{ (k)}}^4 \frac{\delta^2}{\delta x_k^2} = \Delta - \frac{1}{c^2} \frac{\delta^2}{\delta t^2} \dots \dots \dots (11)$$

h. The following potential equations are assumed :

$$\square \Omega = -P, \dots \dots \dots (12)$$

$$\text{Div } \Omega = 0, \dots \dots \dots (13)$$

$$\text{Curl } \Omega = (\mathbf{H}, -i\mathbf{E}) \dots \dots \dots (14)$$

On this basis *for empty space* conceptions are introduced which stand for Lorentz-transformation. But the system is not perfect. Specially the six-dimensional vectors are difficult to fit in a simple way.

5. *Klein's Hint.*

The prominent German geometer, Felix Klein, presents one exception from the general disbelief in the possibilities of the quaternion system. Shortly before his death, while Einstein's theory of relativity was still new, and treated as such by Klein, he interposed a short remark on quaternions as a fit means of rigid transformation of Maxwell's equations.

Introducing the four-dimensional nabla-operator ("Hamiltonian") \diamond as follows :

$$i \frac{\delta}{\delta x_1} + j \frac{\delta}{\delta x_2} + k \frac{\delta}{\delta x_3} + \frac{\delta}{\delta x_4} = \diamond, \dots \dots \dots (15)$$

he points out the possibility of writing the "A-system" of Maxwell's equations for empty space *without any current* in the following way :—

$$\diamond(i(\lambda_{14} + \lambda_{23}) + j(\lambda_{24} + \lambda_{31}) + k(\lambda_{34} + \lambda_{12})) = 0, \dots \dots (16)$$

$$(i(\lambda_{14} - \lambda_{23}) + j(\lambda_{24} - \lambda_{31}) + k(\lambda_{34} - \lambda_{12})) \diamond = 0. \dots \dots (17)$$

The terms λ_{23} , λ_{31} , λ_{12} and λ_{14} , λ_{24} , λ_{34} denote the three-dimensional vector components of \mathbf{H} and \mathbf{E} respectively.

6. *My Plan.*

I shall attempt to concentrate all four Maxwell's equations in one single synthetic quaternion expression

$$\diamond x \psi = 0, \dots \dots \dots (\text{FI})$$

where ψ represents a quaternion of complex members (accordingly eight-dimensional). Eventually (in a later article) it will embrace the complete "B-system." To begin with, simplifications will be made, but still a step further than "Klein's hint" will be taken.

The spirit of the quaternion system as a "group" will be considered, and also the light hereby thrown on Minkowski's potentials.

The expression of all four-dimensional universes, even Lorentz transformation, will be prepared in "quaternion language."

Although I have much material ready on this subject I shall, however, confine this present article to a limited introduction.

7. The Four-dimensional Differentiation.

In the three-dimensional vector analysis the nabla operator ∇ , representing the differentiation $\frac{\delta}{\delta \mathbf{r}}$ with regard to the radius vector, is treated

as an ordinary vector, which may be subject to scalar multiplication with a scalar (gradient) or a vector (divergence). Or a vector multiplication (curl) with a vector may be brought into effect. The four-dimensional nabla operator \Diamond (15) I shall make use of in a corresponding manner.

The radius vector of Minkowski's four-dimensional universe shall be the quaternion :

$$L = \mathbf{r} + l = x\mathbf{i}_1 + y\mathbf{i}_2 + z\mathbf{i}_3 + ict, \quad (18)$$

and correspondingly

$$\Diamond = \mathbf{i}_1 \frac{\delta}{\delta x} + \mathbf{i}_2 \frac{\delta}{\delta y} + \mathbf{i}_3 \frac{\delta}{\delta z} + \frac{\delta}{\delta l}, \quad (19)$$

where (20)

$$\frac{\delta}{\delta l} = \frac{\delta}{ic\delta t} = -\frac{i}{c} \frac{\delta}{\delta t}.$$

Accordingly the four-dimensional differentiation of a quaternion will be written as follows :—

$$\begin{aligned} \Diamond \times O = & \mathbf{i}_1 \left(\frac{\delta O_1}{\delta l} + \frac{\delta \sigma}{\delta x} + \frac{\delta O_3}{\delta y} - \frac{\delta O_2}{\delta z} \right) \\ & + \mathbf{i}_2 \left(\frac{\delta O_2}{\delta l} + \frac{\delta \sigma}{\delta y} + \frac{\delta O_1}{\delta z} - \frac{\delta O_3}{\delta x} \right) \\ & + \mathbf{i}_3 \left(\frac{\delta O_3}{\delta l} + \frac{\delta \sigma}{\delta z} + \frac{\delta O_2}{\delta x} - \frac{\delta O_1}{\delta y} \right) \\ & + \frac{\delta \sigma}{\delta l} - \frac{\delta O_1}{\delta x} - \frac{\delta O_2}{\delta y} - \frac{\delta O_3}{\delta z} \end{aligned}$$

$$=\text{curl } \mathbf{O} + \nabla(\sigma - \mathbf{O}) - \frac{i}{c}(\dot{\sigma} + \dot{\mathbf{O}});$$

$$\frac{\delta}{\delta t}(\sigma + \mathbf{O}) = -\frac{i}{c}(\dot{\sigma} + \dot{\mathbf{O}}). \quad . \quad . \quad . \quad (21)$$

(compare 2). Of course \diamond may be used also as postfactor. Even the analogy of the conjugate quaternion may be introduced.

8. *The Primary Potential ψ .*

The quaternion ψ (F I) may be conceived as a primary potential, from which Maxwell's equations are to be deduced by differentiation in four orthogonal directions. The beautiful "group character" of the quaternion system, according to which addition, multiplication, division, differentiation, etc. of two quaternions always result in a new quaternion (without increase or loss of "dimensions"), together with the extraordinary "closed" character of the system, suggests to me the likelihood that if nature can be expressed in such a system *nothing new can enter and no fundamental character can leave it*, however the equations are treated or manipulated. *Every characteristic feature must be present from the beginning at least as a germ.*

This has induced me to a theory of what I call the "Nucleus of four-dimensional space," hereby meaning a quaternion containing the germs of a mechanical *and* electromagnetic universe, by means of which a certain connexion or "interpretation" between both may be established.

In this article, however, I am only trying to confer the feeling that ψ ought to be so designed that even the most general "B-system" may be deduced in a simple way.

I shall write the general shape as follows :

$$\psi = H - iE = c\mathbf{H} + \kappa - i(c\mathbf{E} + \phi), \quad . \quad . \quad . \quad . \quad . \quad (F II)$$

and suppose that no more members are needed as germs for all electromagnetic conceptions. Accordingly the two scalars κ and ϕ , which have been added to the electromotive and magnetic forces, should contain the "germs" of, for instance, electric current, charge, and magnetic induction according to the idea that nature's laws follow so simple and "closed" a mathematical texture.

Proceeding to develop the four-dimensional differential quotient

$$G = \diamond \times \psi = \diamond \times H - i \diamond \times E = \mathbf{G} + g + i(\mathbf{F} + \gamma) = 0, \quad . \quad . \quad . \quad (22)$$

which must be another (complex) quaternion, and see what happens if I make it equal to zero, I arrive at the following table according to (21) :

	$\mathbf{G}.$	$g.$	$i\mathbf{F}.$	$i\gamma.$
$c \diamond \times \dot{\mathbf{H}} =$	$c \text{ curl } \mathbf{H}$	$-c \nabla \mathbf{H}$	$-i \dot{\mathbf{H}}$	
$+ \diamond \times \dot{\kappa} =$	$+\nabla \kappa$			$-\frac{i}{c} \dot{\kappa}$
$-ic \diamond \times \dot{\mathbf{E}} =$	$-\dot{\mathbf{E}}$		$ic \text{ curl } \mathbf{E}$	$+ic \nabla \mathbf{E}$
$-i \diamond \times \dot{\phi} =$		$-\frac{1}{c} \dot{\phi}$	$-i \nabla \phi$	

. . . (23)

As each of the four components \mathbf{G} , g , \mathbf{F} , γ must disappear individually in order to make the total quaternion equal to zero, I deduce from this:

$$\mathbf{G} = c \text{ curl } \mathbf{H} + \nabla \kappa - \dot{\mathbf{E}} = 0, \quad (24)$$

$$g = -c \nabla \mathbf{H} - \frac{1}{c} \dot{\phi} = 0, \quad (25)$$

$$\mathbf{F} = -\dot{\mathbf{H}} - c \text{ curl } \mathbf{E} - \nabla \phi = 0, \quad (26)$$

$$\gamma = -\frac{1}{c} \dot{\kappa} + c \nabla \mathbf{E} = 0. \quad (27)$$

Now I begin to approach Maxwell's equations. The complete identification depends in the first place on the conception of the scalars κ and ϕ .

9. The Conditions of the First Step: the "A-system."

In order to go somewhat farther than Klein's hint I shall, at least as a first step, try to include the electric current \mathbf{J} . I also have a theory which includes the magnetic induction \mathbf{B} ("the second step") and also the dielectric displacement \mathbf{D} ("the third step"), which, however, must be reserved for some other time, although this "first step" really cannot be rounded off without it. In this article, however, I am endeavouring to stray as little as possible from accustomed and acknowledged paths.

It will easily be seen that the last four equations of the previous paragraph may be identified with the "A-system" of Maxwell's equations subject to the following

"CONDITIONS OF THE FIRST STEP":

$$\nabla \phi = 0, \quad (28)$$

$$\dot{\phi} = 0, \quad (29)$$

$$\nabla \mathbf{v} = -\rho \mathbf{v} = -\mathbf{J}, \quad (30)$$

$$\dot{\kappa} = c^2 \rho, \quad (31)$$

in which case (24) is identified with (A II),

$$(25) \quad , , \quad , , \quad (A \text{ III}),$$

$$(26) \quad , , \quad , , \quad (A \text{ I}),$$

$$(27) \quad , , \quad , , \quad (A \text{ IV}).$$

In quaternion language the first two conditions may be combined into

$$\diamond \times \phi = 0 \quad (F \text{ III})$$

and the two latter into

$$\diamond \times \kappa = -c P_M = -(\mathbf{J} + i\rho c) = -c \diamond \times (\mathbf{H} - i\mathbf{E}), \quad . . . (F \text{ IV})$$

where P_M represents the four-dimensional current mentioned in "d" of paragraph 4.

10. *Minkowski's Potentials.*

Before discussing these results I shall for a moment consider Minkowski's potentials $\Omega_M, \mathbf{A}_M, \phi_M$ (not to be mistaken for my scalar ϕ) as mentioned in "c" of paragraph 4 (compare (5), (6), (7), (8)).

First of all I shall introduce "the conjugate differential operator $\overline{\diamond}$," as follows :

$$\overline{\diamond} = -\mathbf{i}_1 \frac{\delta}{\delta x} - \mathbf{i}_2 \frac{\delta}{\delta y} - \mathbf{i}_3 \frac{\delta}{\delta z} - \frac{i}{c} \frac{\delta}{\delta t} = -\nabla - \frac{i}{c} \frac{\delta}{\delta t}, \quad . . . (32)$$

and its application,

$$\overline{\diamond} \times O = -\text{curl } O + \nabla O - \nabla \sigma - \frac{i}{c} (\dot{\sigma} + \dot{O}). \quad (33)$$

Secondly, I shall introduce special terms for the vector part and the scalar part of the ordinary four-dimensional differential quotient as follows :

$$\diamond \times O = \text{Curl } O - \text{Div } O, \quad (34)$$

defining

$$\text{Curl } O = \text{curl } O + \nabla \sigma - \frac{i}{c} \dot{O}, \quad (F \text{ V})$$

$$\text{Div } O = \nabla O + \frac{i}{c} \dot{\sigma}, \quad (F \text{ VI})$$

and correspondingly for the conjugate,

$$\overline{\diamond} \times O = \bar{\text{Curl}} O - \bar{\text{Div}} O, \quad (34 a)$$

$$\text{Curl } \dot{O} = -\text{curl } \mathbf{O} - \nabla \sigma - \frac{i}{c} \dot{\mathbf{O}}, \quad \dots \quad (\text{F V } a)$$

$$\bar{\text{Div}} \, O = -\nabla \mathbf{O} + \frac{i}{c} \dot{\sigma}. \quad \dots \quad (\text{F VI } a)$$

Thirdly, I shall give Minkowski's potentials the following quaternion shape :

$$-\Omega_M = \mathbf{A}_M + i\phi_M, \quad \dots \quad (35)$$

and deduce

$$\bar{\text{Curl}} \, \Omega_M = \text{curl } \mathbf{A}_M + i\nabla \phi_M + \frac{i}{c} \dot{\mathbf{A}}_M = \mathbf{H}_M - i\mathbf{E}_M, \quad \dots \quad (36)$$

$$\bar{\text{Div}} \, \Omega_M = \nabla \mathbf{A}_M + \frac{1}{c} \dot{\phi}_M. \quad \dots \quad (37)$$

It is evident now how (9), (10), and (14) acquire a much more precise expression in the quaternion language and *how the specially constructed six-dimensional vector of Minkowski's system is entirely absorbed by the quaternion formulas without any necessity for exceptional suppositions.*

The four-dimensional Laplace operator \square enters as follows :

$$\square O = \diamond \times \overline{\diamond} \times O = \overline{\diamond} \times \diamond \times O, \quad \dots \quad (\text{F VII})$$

as may be easily verified.

11. *Have Minkowski's Assumptions been too hasty?*

So far the quaternions have supplied an extraordinarily beautiful means of systematizing laws of nature on the line of Minkowski's ingenious four-dimensional representation. They are even surpassing this, for instance, in the manner of coordinating the six-dimensional conceptions, and also in the indication of a higher mathematical system in the background.

When we attempt to proceed, however, it becomes somewhat more difficult to coordinate Minkowski's assumptions with the quaternion structure of higher order, which I am trying here to build.

To begin with, Minkowski does not take into consideration at all the possibility of the scalars κ and ϕ . Of course it may be said that Minkowski's is a special case, valid for empty space only; still it seems to narrow the view too much, and the possibility of the structure (F IV) of four-dimensional current is lost sight of altogether.

Then Minkowski makes the assumption (7). But *with what right?* From his point of view it may not have seemed preposterous. "The world of the vector analysis is large and seemingly unlimited." So he introduces new conceptions, which he feels quite at liberty to combine

12. *Conclusion.*

My object has been to point out the—in my opinion—too much slighted value of Hamilton's Quaternions as an instrument for treating modern physics. Not only do I consider the difficulties exaggerated, but even that a finer and truer conception may be obtained in this way. The previous paragraph on Minkowski's potentials is an attempt to illustrate what I mean.

A hypothesis is meant to promote scientific investigation until further insight has been acquired, not to postulate a truth. In this light Minkowski's system has been extraordinarily valuable. *But has not Minkowski been too hasty in disclaiming the adaptability of Hamilton's Quaternions to modern physics?* I leave the reader to judge.

If the reader agrees that the point of view presented in this article is worth following up, what has been said here may be considered as an example only, and further developments may follow. For instance, the application to the Lorentz-transformation does not seem very difficult.

References.

- Felix Klein, "Vorlesungen über Nicht-Euklidische Geometrie," 'Vorlesungen über die Entwicklung der Mathematik in 19 Jahrhundert,' Teil ii.
Article by the author: "Näherungslösung zur Ermittlung der wirklichen Spannungsverteilung an konzentriert belasteten Zylinderenden, Ingenieur-Archiv, 1931, Band ii. p. 178.

Kungsgatan 27, Stockholm,
October 18, 1938.

XXXIV. *Notices respecting New Books.*

Applied Thermodynamics. By V. M. FAIRES. [Pp. xvi+374.] (New York : The Macmillan Company, 1938. Price 17s.)

Elementary Thermodynamics. By V. M. FAIRES. [Pp. xii+225.] (New York : The Macmillan Company, 1938. Price 12s.)

THERE is a strong contrast between the typical approaches of practical engineers and of theoretical physicists to the subject of thermodynamics. The engineers, concerned with applications, in dealing with principles, rush in where physicists at least tread warily. The physicists, concerned with principles, often tread so warily that they do not reach applications at all. These two books, by an engineer for engineering students, if they are not free from what the physicist would regard as faults, do possess, in strong measure, the engineering virtues.

The books may appropriately be considered together, as the smaller 'Elementary Thermodynamics' consists in the main of selected portions of the larger 'Applied Thermodynamics' suitable for a shorter course. It has, however, a number of tables and a collection of problems in an appendix, not included in the larger volume, which is intended to be used in conjunction with a separate book of problems.

The first few chapters deal with the general energy equation (introducing the first law), the equations for perfect gases and the simpler processes for gases, and the Carnot and Ericsson cycles (introducing the second law). The treatment is clear and straightforward and generally sound. There is, however, occasional looseness of statement. In the discussion of Charles's Law, for example, it is stated that the determination of the volume of air at two fixed temperatures "provides an absolute scale [of temperature] that depends upon the properties of air. If data for some other gases were available, we should find a slightly different absolute zero point."

Subsequent chapters deal with cycles for internal combustion engines, liquids and vapours, vapour cycles, refrigeration, flow in nozzles, and mixtures of gases and vapours. The larger book covers additional topics. There are chapters on combustion, variable specific heat, and the transfer of heat by conduction and radiation. The performance of internal combustion engines is discussed more fully, and there are much more complete accounts of steam-power plant and of such questions as the flow of steam in nozzles and turbines. The books contain an abundance of excellent diagrams, and also photographic illustrations, both with adequate descriptive captions, and there are many useful tables. Biographical sketches of some of those who have made outstanding contributions to thermodynamics are given as footnotes.

A physics student who picked up these volumes in the commendable desire to know something of what is perhaps the most important field of application of thermodynamics would at once find occasion for surprise in the remarkable system of units which the engineer employs. It might be well, even for the engineering student, if a more complete account were given of units generally, and in particular of that system in which, for example, "J, Joule's constant=

778," and the specific volume of water at 32° F. is 0.016. It is desirable that the student should not regard as essential to the subject peculiarities which arise only from human obstinacy.

These books reveal a great enthusiasm for the subject and a most extensive knowledge of the many branches of engineering to which thermodynamics has been applied. They are frankly text-books, unashamedly pedagogic in manner, and although this will irritate some readers, it will undoubtedly be appreciated by many students. The author, indeed, mentions this in his foreword to the student: "Students who want to become engineers find the course attractive and many of them voluntarily remark that thermodynamics is exceptionally interesting."

E. C. S.

Chemistry of Proteins. By D. JORDAN LLOYD and D. SHORE. Second Edition. (London: Churchill & Co., 1938. Price 21s.)

EVERYONE interested in the nature of protein molecules—pathologists, physiologists, chemists, physicists, and mathematicians—will welcome the appearance of a second edition of this well-known book, which is a gold mine of information regarding the chemical nature of proteins in general. Since the publication of the first edition, the subject of protein chemistry has developed enormously, and in the present edition an attempt has been made to cover in all their important aspects these various new developments. Particular attention has been paid to the vastly important and far reaching work in physical chemistry which has established many important facts about the native proteins, showing them to be globular units with definite molecular weights. The general bearing of these aspects of protein chemistry on biochemistry in general and on physiology and pathology is also considered.

In addition to the very important chapters dealing with the constitutional chemistry of the proteins and their physical chemistry, there are chapters in which questions relating to the architecture of the protein molecule are considered. The possibilities of a geometrical treatment of this important subject are discussed with special reference to the cyclol theory, and some attention is also given to the "numerological" results relating to the numbers of individual amino-acids in different proteins which, if they represent the truth regarding these proteins, indicate mathematical problems whose solution may be helpful in understanding the architecture of protein molecules.

The X-ray studies on proteins are also reported, and it will be seen from the material reviewed in this book that an almost unlimited wealth of material is now available for physical investigations, owing to the immense progress in techniques of isolating and purifying proteins which is so noticeable a feature of recent years.

We commend this book to physicists—and even to mathematicians—in the hope that the fascinating and intriguing picture which it draws of the nature and functions of the protein molecules in living matter will show them how important a field is awaiting further elucidation. Only by combining all these points of view can a solution of the fundamental problems of protein structure be attained.

OBITUARY.

PROF. A. W. PORTER, D.Sc., F.R.S.

It is with very deep regret that we record the death of Prof. A. W. Porter, a valued contributor to this magazine and, at the time of his death, one of the editors.

Alfred William Porter was born on November 12, 1863. He studied at University College, Liverpool, whence he proceeded to University College, London, graduating in the University of London, which then possessed no internal side, in 1890. He was a fellow of University College, and for many years assistant-professor of physics in the College. He was in 1923 appointed to a University chair of physics, retiring therefrom in 1928 with the title of emeritus professor.

Porter's influence on the development of physics was great, if unobtrusive. He filled the position of secretary of the University Board of Studies in physics for many years. He was one of the founder fellows of the Institute of Physics and its first honorary secretary. He was for six years recorder of Section A of the British Association, and in 1928 at the Glasgow meeting, occupied the presidential chair of that section. He was president (1913-14) of the Röntgen Society and (1920-22) of the Faraday Society.

Porter's writings covered a wide range, and illustrate his critical qualities of mind and great erudition. His collaboration, with Carey Foster, in an edition of Joubert's classic volume on Electricity and Magnetism, resulted in a treatise which exercised a not inconsiderable influence on the teaching of that subject. His treatise on Intermediate Mechanics admirably exemplifies his power of developing a subject to a comparatively advanced stage with the help of very elementary mathematics. He was the author of monographs on the method of dimensions and on thermodynamics.

Porter's mind was essentially critical, and he was ever at his best when discussing delicate points concerning the general properties of matter. He contributed to this magazine a valuable series of papers dealing with the shape of the capillary surface under gravity, and his admirable and thoroughly practical analysis of this problem provided results which have earned him the grateful thanks of all those who are concerned with the determination of capillary constants.

Porter made many friends, and a friend once made was never lost. All who knew him well, lament the loss of a kindly companion, a gifted counsellor and a generous and helpful critic.

A. F.

[The Editors do not hold themselves responsible for the views expressed by their correspondents.]

XXXV. *The Calculation of Equilibrium Internuclear Distance for Di-atomic Hydrogen, Hydrides, and Deuterides in Ground and Excited States.*

By C. H. DOUGLAS CLARK, D.Sc., and JOHN L. STOVES, Ph.D.*

[Received September 26, 1938.]

1. *Introduction.*

THE authors have been occupied for some time in making a comparative evaluation of various suggested methods of calculating equilibrium nuclear distance from other band-spectral constants. The comparison has been carried through for non-hydrides ⁽¹⁾, with the general conclusion that the formula proposed by Douglas Clark reproduces the experimental results somewhat better than other suggested relations in these cases. Some preliminary discussion has been given to hydrides ⁽²⁾, but without consideration of electron configurations, a factor which seems from many points of view to be especially important.

Hydrides present a case of great interest from this standpoint, since the quantum states of the assemblage of electrons in the outer system of a hydride di-atom are essentially different from those of the constituent atoms, so that it appears necessary to examine the problem in the light of the electron configurations of the "united-atoms" ⁽³⁾. In the method proposed by Huggins ⁽⁴⁾ it appears that account has been taken of this matter in assigning constants, and we propose to show the analogy existing between his system and ours, and to examine the relative capacity of the two methods to reproduce the experimental data. It appears quite definitely that the constants involved are very sensitive to change in structure, and that in assigning them it is essential to take account not only of the number but also of the nature of the bonding electrons.

2. *Hydride Di-atoms of the KH, LH... Periods.*

In our earlier paper ⁽¹⁾ a system of classifying atoms as s , p , d was suggested, on the basis of the outermost incomplete ground state configurations. This system may be extended to hydride di-atoms by dividing them into two main types, σ and π , with subdivision of the σ type into $s\sigma$, $p\sigma$, and $d\sigma$. All the cases noticed in this communication, including the deuterides considered later, are classified according to this scheme in Table I.

* Communicated by the Authors.

The formulæ here compared are due to Allen and Longair (A) ⁽⁵⁾, Badger (B) ⁽⁶⁾, Huggins (H) ⁽⁴⁾, and Douglas Clark (C) ^{(7) (1)}, as follows:—

$$r_e = [K/(\omega_e \mu^{\frac{1}{2}})]^{\frac{1}{3}}, \quad \text{. (A)}$$

$$r_e = d_{ij} + [5.237 \times 10^6/(\omega_e^2 \mu)]^{\frac{1}{3}}, \quad \text{. (B)}$$

$$r_e = r_{12} - (2.303 \log C)/a, \quad \text{. (H)}$$

$$r_e = [i(k - k')/(\omega_e n^{\frac{1}{2}})]^{\frac{1}{3}}, \quad \text{. (C)}$$

TABLE I.

Classification of Hydrides into Types and Sub-types.

Main type.	Sub-type.	Group number.	Hydride di-atoms and di-atom ions.
$\sigma \dots\dots$	$s\sigma$	1	HH ⁺
		2	HH, HD, DD, LiH, LiD, NaH, CuH, AgH, AuH, BeH ⁺ , BeD ⁺ , MgH ⁺ , ZnH ⁺ , CdH ⁺ , HgH ⁺ .
	$p\sigma$	3	BeH, MgH, CdH, HgH.
		4	AlH, AlD.
	$d\sigma$	3	CaH.
$\pi \dots\dots$	$p\pi$	6	BiH.
		7	OH, ClH ⁺ .
		8	FH, ClH, ClD, BrH.

where

(a) K, d_{ij}, k, r_{12} are period constants, the last-named having an "additive" character ;

(b) a is a group constant, which also varies somewhat with period ;

(c) k' is a correction for di-atom ions ;

(d) μ is the reduced mass, equal to $1.649M$, in 10^{-24} gram units, where $M = A_1 A_2 / (A_1 + A_2)$, and A_1, A_2 are the atomic weights of the constituent atoms of a di-atom on the 0=16 scale ;

(e) n is the group number, equal to the total number of valency electrons of the atoms concerned ;

(f) i is a correction for isotopes, approximately equal to unity, except for HD ($i = \sqrt{3}/2$) and DD ($i = 1/\sqrt{2}$), and all other deuterides ;

TABLE II.
Constants used for Hydride Neutral and Ion Periods.

Period.	A. K.	B. d_{ij}	H.					C.		
			$a.$		r_{12}'			$k-k'$		$p\pi.$
			$\sigma.$	$\pi.$	$\sigma.$	$\pi.$	$\pi.$	$s\sigma.$	$p\sigma.$	
KH	5,260	0.335	3.5	6.0	1.59	1.11	1.11	7,440	8,550	8,870
LH	9,780	0.585	3.5	6.0	1.81	1.34	1.34	12,100	13,830	(17,410)
MH	12,450	0.650	3.5	..	1.92	17,280 ($d\sigma$)
$\bar{M}H$	12,450	0.650	4.5	6.0	1.55	1.49	1.49	9,470	..	(21,010)
$\bar{N}H$	9,770	0.640	4.5	..	1.66	10,410	13,560	..
$\bar{O}H$	11,150	0.697	4.5	6.0	1.66	1.63	1.63	11,400	13,430	23,750
KH^+	5,260	0.335	3.5	6.0	1.59	1.11	1.11	7,870
LH^+	9,780	0.585	3.5	6.0	1.81	1.34	1.34	11,910	..	(15,120)
$\bar{M}H^+$	12,450	0.650	4.5	..	1.55	9,490
$\bar{N}H^+$	9,770	0.640	4.5	..	1.66	11,490
$\bar{O}H^+$	11,150	0.697	4.5	..	1.66	11,350

TABLE III.—Calculated Results in

Period.	Di-atom and type.	State.	Group number <i>n</i> .	μ $\times 10^{24}$ gr.	ω_e cm. ⁻¹ .	$\omega_e x_e$ cm. ⁻¹ .	r_e (expt.) Å.U.
KH	LiH <i>sσ</i>	A ¹ Σ ⁺	2	1.453	279.8	-9.56	2.586 <i>c</i>
		X ¹ Σ ⁺	1406.1 <i>a</i>	22.73	1.593 <i>c</i>
	BeH <i>pσ</i>	A ² Π _{reg.}	3	1.495	2087.6	39.85	1.330
		X ² Σ ⁺	2058.5 <i>a</i>	35.5	1.340
	OH <i>pπ</i>	B ² Σ ⁺	7	1.563	3182.5	97.7	1.009
		² Σ ⁺	3176.7 <i>b</i>	92.80	1.015
		² Π	3728.0 <i>b</i>	78.15	0.974
	FH <i>pπ</i>	¹ Σ	8	1.578	4123.1	90.8	0.915
LH.....	NaH <i>sσ</i>	A ¹ Σ ⁺	2	1.592	335.2	-4.42	3.03
		X ¹ Σ ⁺	1170.8 <i>a</i>	18.9	1.88
	MgH <i>pσ</i>	A ² Π _{reg.}	3	1.595	1603.5	34.8	1.68
		X ² Σ ⁺	1493.5 <i>a</i>	31.3	1.73
	AlH <i>pσ</i>	A ¹ Σ ⁺	4	1.602	1682.6 <i>a</i>	29.15	1.643
	ClH <i>pπ</i>	¹ Σ	8	1.615	2989.7 <i>a</i>	51.65	1.272
MH	CaH <i>dσ</i>	C ² Σ ⁺	3	1.621	1448	26	1.85
		D ² Σ ⁺	1147.5	32.3	2.61
		B ² Σ ⁺	1285	20	1.95
		A ² Π _{reg.}	1333	20	1.99
		X ² Σ ⁺	1316.7 <i>a</i>	19.2	2.01
MH	CuH <i>sσ</i>	B ¹ Σ	2	1.636	1794.5	115.9	1.610
		A ¹ Σ	1701.7	45.15	1.569
		X ¹ Σ	1939.9 <i>a</i>	37.15	1.460
	BrH <i>pπ</i>	¹ Σ	8	1.643	2649.7 <i>a</i>	44	1.410
NH ...	AgH <i>sσ</i>	A ¹ Σ	2	1.646	1663.6	87.0	1.638
		X ¹ Σ	1760.0 <i>a</i>	34.05	1.614
	CdH <i>pσ</i>	B ² Σ ⁺	3	1.647	874.1	18.83	2.377
		A ² Π _{reg.}	1749.8	34.65	1.65
		X ² Σ ⁺	1430.7 <i>a</i>	46.3	1.76
OH	AuH <i>sσ</i>	A ¹ Σ	2	1.653	1690	72	1.67
		X ¹ Σ	2302 <i>a</i>	42	1.53
	HgH <i>pσ</i>	A ² Π _{reg.}	3	1.654	2063.2	39.8	1.577
		X ² Σ ⁺	1432.7	108.8	1.729
	BiH <i>pπ</i>	B ¹ Σ	6	1.654	1695.5	(17.5)	1.775
		A ¹ Π	1716.9	23.87	1.784
		X ¹ Σ	1678.2 <i>a</i>	21.23	1.804

a. Ground state.*b.* See reference no. 8.*c.* Constants for LiD.

the Periods of Hydride Di-atoms.

r_e (calc.).				r_e percentage errors.			
A.	B.	H.	C.	A.	B.	H.	C.
2.499	3.918	2.682 <i>d</i>	2.724	— 3.36	+51.51	+ 3.71	+ 2.82
1.459	1.557	1.631	1.591	— 8.41	— 2.26	+ 2.39	— 2.57
1.603	1.265	1.313	1.332	+20.52	— 4.89	— 1.28	+ 0.15
1.610	1.274	1.343	1.339	+20.15	— 4.97	+ 0.22	— 0.07
1.097	1.026	0.995	1.017	+ 8.73	+ 1.69	— 1.39	+ 0.79
1.098	1.027	1.001	1.018	+ 8.18	+ 1.18	— 1.38	+ 0.29
1.041	0.957	0.961	0.965	+ 6.88	— 1.72	— 1.34	— 0.92
1.006	0.915	0.915	0.913	+ 9.95	0.00	0.00	— 0.22
2.849	3.667	2.752 <i>d</i>	2.852	— 5.97	+21.02	— 9.18	— 2.84
1.878	1.924	1.938	1.880	— 0.11	+ 1.45	+ 3.08	+ 3.19
1.690	1.670	1.679	1.708	+ 0.60	— 0.60	+ 0.00	+ 1.67
1.731	1.722	1.739	1.748	+ 0.06	— 4.63	+ 0.52	+ 1.04
1.663	1.634	1.673 <i>e</i>	1.602	+ 1.22	— 0.55	+ 1.83	— 2.51
1.371	1.298	1.282 <i>f</i>	(1.272)	+ 7.78	+ 2.04	+ 0.79	(0.00)
1.890	1.805	1.885	1.903	+ 2.16	— 2.43	+ 1.89	+ 2.93
2.042	1.999	1.986	2.056	(—21.76)	(—23.41)	(—23.91)	(—21.22)
1.965	1.901	1.983	1.980	+ 0.77	— 2.51	+ 1.69	+ 1.54
1.943	1.870	1.962	1.956	— 2.36	— 6.03	— 1.41	— 2.21
1.951	1.880	1.973	1.963	— 2.94	— 6.47	— 1.84	— 2.34
1.757	1.648	1.286	1.551	+ 9.13	+ 2.36	—20.12	— 3.67
1.774	1.684	1.556	1.579	+13.07	+ 7.33	— 0.83	+ 0.80
1.712	1.598	1.516	1.511	+ 1.73	+ 9.45	+ 3.84	+ 3.49
1.542	1.419	1.477 <i>d</i>	(1.410)	+ 9.36	+ 0.64	+ 4.75	(0.00)
1.661	1.688	1.553	1.642	+ 1.40	+ 3.05	— 5.19	+ 0.24
1.629	1.649	1.674	1.612	+ 0.93	+ 2.17	+ 3.72	— 0.12
2.058	2.249	2.019 <i>d</i>	2.077	(—13.42)	(— 5.39)	(—15.06)	(—12.62)
1.633	1.653	1.676	1.647	— 1.03	+ 0.18	+ 1.58	— 0.18
1.746	1.798	1.738	1.762	— 0.98	+ 2.66	— 1.25	+ 0.14
1.725	1.632	1.593	1.683	+ 3.29	— 2.28	— 4.61	+ 0.78
1.556	1.540	1.535	1.518	+ 1.70	+ 6.54	+ 0.33	— 0.78
1.614	1.603	1.692	1.555	+ 2.35	+ 1.65	+ 7.29	— 1.39
1.823	1.852	1.528	1.756	+ 5.44	+ 7.12	—11.62	+ 1.54
1.723	1.730	1.788 <i>d</i>	1.788	— 2.93	— 2.54	+ 0.73	+ 0.73
1.716	1.721	1.778 <i>d</i>	1.780	— 3.81	— 3.53	— 0.34	— 0.22
1.729	1.729	1.788 <i>d</i>	1.795	— 4.16	— 4.16	— 0.89	— 0.50

d. a' negative.

e. Same constants as for NaH.

f. a' zero.

TABLE IV.
Calculated Results in the Periods of Hydride Di-atom Ions.

Period.	Di-atom and type.	State.	Group number n .	μ $\times 10^{24}$ gr.	ω_e , cm. ⁻¹ .	$\omega_e x_e$, cm. ⁻¹ .	r_e (expt.), A.U.	r_e (calc.).			r_e percentage errors.				
								A.	B.	H.	A.	B.	H.	C.	
KH ⁺ . . .	{ BeH ⁺ σ }	{ A ¹ Σ ⁺ X ¹ Σ ⁺ }	2	1.495	1476.1	14.8	1.61	1.429	1.506	1.627 <i>c</i>	1.556	-11.24	-6.46	+1.06	-3.36
			2220.0 <i>a</i>	39.8	1.313	1.247	1.227	1.374	1.358	-5.03	-6.55	+4.65	+3.43
LH ⁺ . . .	{ MgH ⁺ σ }	{ A ¹ Σ ⁺ X ¹ Σ ⁺ }	2	1.595	1139.0	9.2	2.010	1.894	1.947	1.997 <i>c</i>	1.948	-5.77	-3.13	-0.65	-2.59
			1695.3 <i>a</i>	30.2	1.653	1.659	1.630	1.671	1.706	+0.36	-1.39	+1.09	+3.28
	{ ClH ⁺ π π }	² Σ ⁺	7	1.615	1603	38	1.528	1.687	1.666	1.503 <i>c</i>	(1.528)	+10.41	+9.03	-1.64	(0.00)
MH ⁺ . . .	{ ZnH ⁺ σ }	{ A ¹ Σ ⁺ X ¹ Σ ⁺ }	2	1.637	1365 <i>b</i>	15	1.71	1.924	1.848	1.720 <i>c</i>	1.700	+12.51	+8.07	+0.58	-0.59
			1915 <i>ab</i>	39	1.51	1.719	1.605	1.562	1.518	+13.84	+6.29	+3.44	+0.53
NH ⁺ . . .	{ CdH ⁺ σ }	{ A ¹ Σ ⁺ X ¹ Σ ⁺ }	2	1.647	1250 <i>b</i>	8.6	1.861	1.826	1.906	1.881 <i>c</i>	1.865	-1.88	+2.42	+1.08	+0.21
			1773 <i>ab</i>	37.3	1.664	1.625	1.643	1.702	1.661	-2.34	-1.26	+2.28	-0.18
OH ⁺ . . .	{ HgH ⁺ σ }	{ A ¹ Σ ⁺ X ¹ Σ ⁺ }	2	1.654	1647	69.5	1.69	1.740	1.750	1.786	1.695	+2.96	+3.55	+5.68	+0.30
			2016 <i>a</i>	40.8	1.59	1.627	1.617	1.651	1.585	+2.32	+1.70	+3.84	-0.31

a, Ground state, *b*, Band-head measured, *c*, *a'* negative.

a. Ground state.

b. Band-head measured.

c. a' negative.

(g) C , a' are given respectively by the relations

$$C = 5.85 \times 10^{-6} M \omega_e^2 / (a^2 - aa'), \quad (H\ 1)$$

$$a' = (2.0625 a^2 + 0.7154 M \omega_e x_e)^{\frac{1}{2}} - 1.75 a; \quad . . . (H\ 2)$$

(h) r_e (Å.U.), ω_e and $x_e \omega_e$ (cm.⁻¹) have their usual band-spectral significance.

Of the above four relations, (A), (B), (C), and (H), the first three are empirical, whilst the last has theoretical background but is partly empirical in application. We are not here specially concerned with the derivation of the (H) relation: suffice it to say that it is based upon earlier work of Born and others on alkali halide crystals, and that it depends upon a modified Morse exponential potential energy function and well-known relations between the coefficients of the expansion in powers of $(r - r_e)$.

Huggins has found that it is broadly necessary to make assignments of a and r_{12} for hydride di-atoms having $n=2, 3$, or 4 , different from those for hydrides having $n=6, 7$, or 8 valency electrons. This suggests a structural influence, which is closely analogous to that we have employed for non-hydrides. In the earlier communication concerning hydrides ⁽²⁾ certain assumptions were made in the LH period to reconcile calculation with experiment, whereby an "effective group number" was introduced in certain cases. It is now realized that a more satisfactory way of dealing with the problem involves an allowance for the various electron configurations. This has proved valuable for non-hydrides, and when applied to hydrides brings our method into structural relation with that of Huggins.

Table II. (p. 391) shows the assigned constants used for calculating results in Table III. (pp. 392, 393). It appears that Huggins has essentially allowed for the main σ and π types of Table I. (p. 390) in assigning his constants, whilst we have taken account of sub-types as well. It should be said that Huggins has queried some of his assignments, but we have taken them as a provisional basis for calculation in the absence of other numbers. The A constants in Table II. have been chosen as far as possible to suit requirements in the KH, LH, . . . periods, the numbers being transferred unchanged to "barred" periods $\bar{M}H$, $\bar{N}H$. . . and to periods of ions KH^+ , LH^+ The C formula is the only one which allows for electron configurations as well as for ionized states. This very detailed method of allotment has some disadvantages, however, in the present state of the experimental evidence, for in certain cases (ClH, ClH^+ , and BrH) the constants have been necessarily taken from one electron state only. These are shown in parentheses in Table II., and the results are not included in the later calculation of mean errors.

Table III. shows the results of calculation for hydride di-atoms, and Table IV. (p. 394) for hydride di-atom ions. In certain cases, the percentage errors are bracketed, these numbers not being used in calculation of mean errors. (For example, the results suggest experimental error in r_e for $\text{CaH}(\text{D } ^2\Sigma^+)$ and $\text{CdH}(\text{B } ^2\Sigma^+)$.)

TABLE V.
Average Percentage Errors in Hydride Periods (\pm per cent.).

Table no.	Period.	A.	B.	H.	No. of states ABH.	C.	No. of states C.
III. ...	Di-atoms						
	KH	10.77	8.53	3.10	8	0.98	8
	LH	2.62	5.05	2.57	6	2.25	5
	MH	2.06	4.36	1.71	4	2.23	4
	$\overline{\text{MH}}$	8.32	4.95	7.39	4	1.99	3
	$\overline{\text{NH}}$	1.09	2.02	2.94	4	0.17	4
	$\overline{\text{OH}}$	3.38	3.97	3.69	7	0.84	7
IV. ..	Di-atom ions						
	KH^+ to $\overline{\text{OH}}^+$	6.24	4.53	2.36	11	1.48	10
Mean errors.....		4.92	4.77	3.39	7 (period sets)	1.42	7 (period sets)
"Weighted" mean errors..		5.46	5.04	3.18	44	1.40	41

All the calculations have been carefully checked, the accuracy being that of four-figure logarithms, in accordance with earlier procedure. It may be said that the (H) equation involves more tedious calculation as compared with the (A), (B), and (C) relations.

The mean percentage errors (regardless of sign) are in Table V.

The calculation of "weighted" means takes account of the number of states involved and gives the mean errors for all the recorded numbers by each method. It is observed that the (C) formula gives the best results when allowance is made for electron configurations in the way suggested, the mean error in calculating r_e being ± 1.4 per cent., which may well be within the limits of error in assignments of experimental values of r_e .

3. *The Di-atoms HH, HD, and DD : the Isotope Effect.*

Table VI. gives observed and estimated values of constants for ground and excited states of HH, HD, and DD. The ω_e 's, r_e 's, and r_0 's for HH are experimental numbers. The corresponding ω_e 's of HD and DD are calculated from those of HH by multiplying by the i factors, $\sqrt{\mu_{\text{HH}}/\mu_{\text{HD}}} = 0.866$ and $\sqrt{\mu_{\text{HH}}/\mu_{\text{DD}}} = 0.707$, respectively. Good agreement with experiment is obtained in all cases where comparison is possible. The r_e 's (and r_0 's) of HD and DD are obtained, where no experimental numbers are available, by reducing the corresponding values for HH by about 0.15 and 0.3 per cent. respectively. This seems a reasonable procedure, as evidenced by the following experimental r_0 's: State $1s\sigma\ 3p\pi\ ^3\Pi_u$, HH 1.059, HD 1.058, DD 1.056; $1s\sigma\ 2s\sigma\ ^3\Sigma^+$, HH 0.999, HD 0.997, DD 0.996. A self-consistent scheme appears thus to be obtained.

It may be noted that r_e may be obtained from r_0 by means of the relation $1/r_e^2 = (1/r_0^2) + (\alpha\mu)/53.32^*$. Thus, for the ground state of HH, $r_0 = 0.749$, $\alpha = 2.9$, $\mu = 0.831$, whence $r_e = 0.740$. In cases where α is unknown it may be approximated by putting B_0 for B_e and r_0 for r_e in the relation due to Huggins

$$\alpha = (2B_e^2/\omega_e) [r_e(a+a') - 3], \quad \dots \dots \dots (\text{H } 3)$$

obtaining r_e as before when α is found.

Considering the case of HH, it appears that no formula advanced up to the present time very accurately represents the relation between ω_e and r_e for the ground as well as excited states. This may be partly due to the fact that the ground state of HH is unique in having two 1-quantum electrons: $1s\sigma^2$. We have attempted to overcome this difficulty by using three different constants for the (C) formula for the ground singlet state, remaining singlet states, and the triplet states. Similar considerations apply to HD and DD. The discrepancy, using our formula for the ground state of HH, was noted by Sandeman⁽⁹⁾.

The (C) formula is the only one which so far does not compensate for the isotope effect: the other formulæ, (A), (B), and (H) achieve this, because in different ways they each involve ω_e in inverse relation to $\mu^{\frac{1}{2}}$. (Since ω_e is given by $\omega_e = (1/2\pi c)(k_e/\mu)^{\frac{1}{2}}$, where k_e is the "bond constant," which may be presumed not to vary very much between corresponding states of isotopic forms, we have ω_e proportional to $\mu^{-\frac{1}{2}}$.) The isotope effect, in general, does not appear to matter, so far as the available evidence goes, in connexion with the (C) formula for non-hydrides, since the mean error appears to lie within the probable experimental error. This has been shown to be the case for the isotopic di-atoms ^{10}BO and

* Obtained from the relations $B_0 = B_e - 0.5\alpha$, $B_0 = 27.66/(\mu r_0^2)$, and $B_e = 27.66/(\mu r_e^2)$, in the units chosen.

TABLE VI.
Experimental and Estimated Constants for HH, HD, and DD.

State.	HH.			HD.			DD.		
	ω_e	r_e	r_0	ω_e (calc.).	ω_e (expt.).	r_e	r_0	ω_e (calc.).	ω_e (expt.).
$1s\sigma\ 6p\pi\ ^3\Pi_u\ \dots\dots$	2293	1.052	1.065	1986	..	1.051 <i>a</i>	1.064 <i>a</i>	1622	..
								1.049 <i>a</i>	1.062 <i>a</i>
$1s\sigma\ 5p\pi\ ^3\Pi_u\ \dots\dots$	2309	1.052	1.065	1995	..	1.051 <i>a</i>	1.064 <i>a</i>	1629	..
								1.049 <i>a</i>	1.062 <i>a</i>
$1s\sigma\ 4p\pi\ ^3\Pi_u\ \dots\dots$	2336.5	1.051	1.065	2024	..	1.050 <i>a</i>	1.064 <i>a</i>	1653	..
								1.048 <i>a</i>	1.062 <i>a</i>
$1s\sigma\ 3p\pi\ ^3\Pi_u\ \dots\dots$	2372.5	1.035	1.059	2055	2054.6	1.034	1.058	1678	1678.2
								1.032 <i>a</i>	1.056
$1s\sigma\ 3p\sigma\ ^3\Sigma_g^+\ \dots\dots$	2197	1.103	1.120	1903	..	1.102 <i>a</i>	1.119 <i>a</i>	1554	..
								1.100 <i>a</i>	1.117 <i>a</i>
$1s\sigma\ 2s\sigma\ ^3\Sigma_g^+\ \dots\dots$	2664.6	0.96	0.999	2309	2308.4	0.96 <i>a</i>	0.997	1886	1885.8
								0.96 <i>a</i>	0.996
$1s\sigma\ 2p\pi\ ^3\Pi_u\ \dots\dots$	2463	0.986	1.047	2133	..	0.985 <i>a</i>	1.046 <i>a</i>	1742	..
								0.984 <i>a</i>	1.044 <i>a</i>
$1s\sigma\ 3s\sigma\ ^1\Sigma_g^+\ \dots\dots$	2538	1.04	1.06	2199	..	1.04 <i>a</i>	1.06 <i>a</i>	1796	..
								1.04 <i>a</i>	1.06 <i>a</i>
$1s\sigma\ 3d\pi\ ^1\Pi_g\ \dots\dots$	2252.9	1.07	1.09	1952	..	1.07 <i>a</i>	1.09 <i>a</i>	1594	..
								1.07 <i>a</i>	1.09 <i>a</i>
$1s\sigma\ 3d\sigma\ ^1\Sigma_g^+\ \dots\dots$	2455.5	1.05	1.08	2127	..	1.05 <i>a</i>	1.08 <i>a</i>	1738	..
								1.05 <i>a</i>	1.08 <i>a</i>
$1s\sigma\ 2p\sigma\ ^1\Sigma_u^+\ \dots\dots$	1357.3	1.276	1.308	1176	1175.7	1.275 <i>a</i>	1.306	960.3	..
								1.274 <i>a</i>	1.305 <i>a</i>
$1s\sigma^2\ ^1\Sigma_g^+\ \dots\dots$	4403	0.740	0.749	3815	3826.0	0.739	0.748	3116	3115
								0.738 <i>a</i>	0.747 <i>a</i>

a. Estimated values.

$^{11}\text{BO}^{(10)}$. The (C) formula taking account of isotopes for non-ionized states assumes the form $\omega_e r_e^3 \sqrt{n} = ki$, where i is an isotopic factor nearly equal to unity in many cases. For HD and DD, however, the isotopic factor must be taken into account by multiplying the corresponding empirical constants for HH by the factors $i(\text{HD}) = \sqrt{3/2} = 0.866$, and $i(\text{DD}) = 1/\sqrt{2} = 0.707$. In this way the scheme of constants shown in Table VII. is obtained.

TABLE VII.
Constants used for HH, HD, and DD.

Di-atom.	K.	B.	H.		C.			
			$a.$	$r_{12}.$	Ground state.	Other singlet states.	Triplet states.	
					$1s\sigma^2\ ^1\Sigma_g^+$	$1s\sigma 2p\sigma\ ^1\Sigma_u^+$	$1s\sigma 2p\pi^3\ ^3\Pi_u$	Lowest stable state.
HH ..	2.400	0.025	6.0	0.90	2,520	3,980	3,710	
HD ..					2,180	3,450	3,340	
DD ..					1,780	2,820	2,625	

TABLE VIII.
Comparison of Empirical and Calculated k 's for Deuterides.

Di-atom.	Type.	$k.$	X.	$i.$	Di-atom.	Mean k (empirical).	$ki.$
LiH ...	$s\sigma$	7,440	6.94	0.750	LiD	5,254	5,582
BeH ⁺ ..	$s\sigma$	7,870	9.02	0.742	BeD ⁺	5,805	5,837
AlH ...	$p\sigma$	13,830	26.97	0.720	AlD	10,730	9,954
ClH ...	$p\pi$	17,410	35.46	0.717	ClD	12,480	12,480

In general, where necessary, the value of the isotopic factor for k in proceeding from X_1Y_1 to X_2Y_2 may be taken as $i = \sqrt{\mu_{X_1Y_1}/\mu_{X_2Y}} = \sqrt{X_1Y_1(X_2+Y_2)/[X_2Y_2(X_1+Y_1)]}$, where X_1, Y_1, X_2, Y_2 are the atomic weights concerned. In the particular case of XH to XD this reduces to $i = \sqrt{(X+2)/(2X+2)}$, where X represents the atomic weight of the

non-hydrogenic atom *. We now apply this procedure to the deuterides LiD, BeD⁺, AlD, and ClD, for which data are recorded in Table VIII. The *k*'s for hydrides are as in Table II.

TABLE IX.—Calculated Results in

Period.	Di-atom and type.	State.	Group number <i>n</i> .	μ $\times 10^{24}$ gr.	ω_e cm. ⁻¹ .	$\omega_e x_2$ cm. ⁻¹ .	r_e (expt.). Å.U.	
HH ...	HH $1s\sigma\ 2p\pi$	$1s\sigma$ {	$6p\pi\ ^3\Pi_u$	2	0.831	2293	57	1.052
			$5p\pi\ ^3\Pi_u$	2309	58	1.025
			$4p\pi\ ^3\Pi_u$	2336.5	60	1.051
			$3p\pi\ ^3\Pi_u$	2372.5	65.0	1.035
			$3p\sigma\ ^3\Sigma^+$	2197	67	1.103
			$2s\sigma\ ^3\Sigma_g^+$	2664.6	70.6	0.986
			$2p\pi\ ^3\Pi_u$	2463	73	0.96
	HH $1s\sigma\ 2p\sigma$	$1s\sigma$ {	$3s\sigma\ ^1\Sigma_g^+$	2	0.831	2538	124	1.04
			$3d\pi\ ^1\Pi_g$	2252.9	77.8	1.07
			$3d\sigma\ ^1\Sigma_g^+$	2455.5	114.9	1.05
			$2p\sigma\ ^1\Sigma_u^+$	1357.3	19.54	1.276
	HH $1s\sigma^2$	$1s\sigma^2\ ^1\Sigma_g^+$	2	0.831	4403 <i>a</i>	120.5	0.740	
	HD $1s\sigma\ 2p\pi$	$1s\sigma$ {	$3p\pi\ ^3\Pi_u$	2	1.108	2054.6	49.74	1.034
			$2s\sigma\ ^3\Sigma_g^+$	2308.4	53.77	0.985
	HD $1s\sigma\ 2p\sigma$	$1s\sigma$ {	$2p\pi\ ^1\Pi_u$	2	1.108	2140.3	66.72	1.05
			$2p\sigma\ ^1\Sigma_u^+$	1175.7	14.98	1.275
	HD $1s\sigma^2$	$1s\sigma^2\ ^1\Sigma_g^+$	2	1.108	3826.0 <i>a</i>	90.4	0.739	
	DD $1s\sigma\ 2p\pi$	$1s\sigma$ {	$3p\pi\ ^3\Pi_u$	2	1.660	1678.2	32.94	1.032
			$2s\sigma\ ^3\Sigma_g^+$	1885.8	35.96	0.984
	DD $1s\sigma^2$	$1s\sigma^2$	2	1.660	3115 <i>a</i>	60.3	0.738	
KD ...	LiD $s\sigma$	A $^1\Sigma^+$	2	2.573	183.1	-12.74	2.586	
		X $^1\Sigma^+$	1055.1 <i>a</i>	13.23	1.593	
KD+ ..	BeD ⁺ $s\sigma$	A $^1\Sigma^+$	2	2.713	1096.4	8.5	1.602	
		X $^1\Sigma^+$	1647.6 <i>a</i>	21.9	1.31	
LD ...	AlD $p\sigma$	A $^1\Sigma$	4	3.089	1212.0 <i>a</i>	15.15	1.642	
	ClD $p\pi$	$^1\Sigma$	8	3.142	2144 <i>a</i>	26.59	1.272	

a. Ground state.

* The atomic weights of H and D are taken sufficiently accurately for this purpose as 1 and 2 respectively; the more accurate values are 1.0077 and 2.0135.

It is observed that the empirical k 's and the values of ki agree fairly well: the agreement is not exact, except for ClH, for various reasons.

the HH and Deuteride Periods.

r_e (calc.).				r_e percentage errors.			
A.	B.	H.	C.	A.	B.	H.	C.
1.047	1.087	1.060 <i>b</i>	1.046	— 0.47	+ 3.33	+ 0.76	— 0.57
1.046	1.084	1.058 <i>b</i>	1.044	— 0.57	+ 3.04	+ 0.57	— 0.76
1.041	1.074	1.053 <i>b</i>	1.040	— 0.95	+ 2.49	+ 0.19	— 1.05
1.036	1.064	1.045 <i>b</i>	1.034	+ 0.10	+ 2.80	+ 0.97	— 0.09
1.062	1.118	1.070 <i>b</i>	1.061	— 3.72	+ 1.36	— 2.99	— 3.81
0.996	0.986	1.004 <i>b</i>	0.995	+ 1.01	0.00	+ 1.83	+ 0.91
1.023	1.038	1.029 <i>b</i>	1.021	(+ 6.56)	(+ 8.13)	(+ 7.18)	(+ 6.35)
1.012	1.018	0.995	1.035	— 2.69	— 2.12	— 4.33	— 0.48
1.054	1.100	1.057 <i>b</i>	1.077	— 1.50	+ 2.80	— 1.21	+ 0.65
1.105	1.040	1.010	1.047	+ 5.24	— 0.95	— 3.81	— 0.29
1.247	1.532	1.252 <i>b</i>	1.275	— 2.27	+ 20.05	— 1.88	— 0.08
0.843	0.713	0.813	(0.740)	(+ 13.90)	(— 3.65)	(+ 9.87)	(0.00)
1.035	1.063	1.045 <i>b</i>	1.034	+ 0.10	+ 2.81	— 1.23	0.00
0.996	0.986	1.004 <i>b</i>	0.995	+ 1.12	+ 0.10	+ 0.70	+ 1.02
1.022	1.036	1.004 <i>b</i>	1.045	— 2.67	— 2.29	— 3.65	— 0.48
1.247	1.531	1.252 <i>b</i>	1.275	— 2.20	+ 20.07	— 4.14	0.00
0.841	0.712	0.812	(0.739)	(+ 13.80)	(— 3.66)	(+ 9.88)	(0.00)
1.036	1.064	1.045 <i>b</i>	1.034	— 0.39	+ 3.10	— 1.04	+ 0.19
0.996	0.986	1.003 <i>b</i>	0.995	+ 1.22	+ 0.20	+ 0.70	+ 1.12
0.842	0.712	0.813	(0.738)	(+ 14.09)	(— 3.52)	+ 10.16	(0.00)
2.616	4.265	1.496 <i>b</i>	2.728	+ 1.16	+ 64.94	— 42.15	+ 5.49
1.460	1.558	1.626	1.522	— 8.35	— 2.19	+ 2.07	— 4.46
1.428	1.506	1.679 <i>b</i>	1.552	— 10.86	— 5.99	+ 4.81	— 3.12
1.257	1.228	1.275	1.356	— 4.05	— 6.25	— 2.67	+ 3.51
1.662	1.634	1.672	1.642	+ 1.22	— 0.49	+ 1.83	(0.00)
1.371	1.298	1.281	1.272	+ 7.78	+ 2.04	+ 0.71	(0.00)

b. a' negative.

In the excited state of LiD ($A\ ^1\Sigma^+$) the value of ω_e would be expected to be $\omega_e(\text{LiH}) \times i = 279.8 \times 0.75 = 209.9$, whereas the recorded $\omega_e(\text{LiD})$ is

183.1 cm.⁻¹. In the case of AlD the discrepancies are due to the different states used in estimating $k(\text{AlH})$, since the data for MgH were included (also of LH $\rho\sigma$ type). Actually, except in the case of HH, multiplication of k by the isotopic factor i for a given di-atom is not an entirely suitable procedure, the reason being that k properly refers to a period type and not to an individual di-atom. With more data it might be possible to estimate the ki 's of deuteride period types more accurately, by taking an average value of the atomic weight X for the actual cases included. On the whole, until further experimental results are available it seems to be best to employ the empirical k 's for deuterides other than HD and DD. We therefore use the constants of Table VII. for HH, HD, and DD, the empirical k 's of Table VIII. for the deuteride di-atoms using the (C)

TABLE X.
Average Percentage Errors in HH and Deuteride Periods.

Periods.	Di-atoms.	A.	B.	H.	No. of states ABH.	C.	No. of states C.
HH.....	HH	1.85	3.89	1.85	10	0.87	10
	HD	1.52	6.32	2.43	4	0.38	4
	DD	0.81	1.65	0.87	2	0.65	2
KD, KD ⁺ , LD	Deuterides	5.57	13.65	9.04	6	4.15	4
Mean errors.....		2.44	6.38	3.55	4 (period sets)	1.51	4 (period sets)
"Weighted" mean errors .		2.71	6.79	3.83	22	1.40	20

formula, and the same constants for deuterides as for the corresponding hydrides in the (A), (B), and (H) equations *. Table IX. has been obtained in this way.

The errors enclosed in parentheses in Table IX. have not been included in calculating mean errors. One state of HH ($1s\sigma 2p\pi \ ^3\Pi_u$) is omitted, since all methods give large positive errors (6 to 8 per cent.), and the experimental number used is probably too low. The "weighted" mean percentage errors again give the lowest error (1.40 per cent.) for the C formula.

It thus appears that the formula proposed by us, when full allowance

* It may be noticed that Huggins' r_{12} is not truly additive for HH: thus the assignment for H is $r_{12}=0.38$ and for HH, when $a=6$, $r_{12}=0.90$, unequal to $2 \times 0.38=0.76$.

is made for structural considerations, is capable of reproducing the experimental results with considerable precision. The mean error for the excited states of HH, HD, and DD is under 1 per cent. It has also been noted that our formula is more convenient in application than Huggins' relation.

The values of the various experimental constants used have been largely taken from well-known summaries by Jevons⁽¹¹⁾ and Miss Sponer⁽¹²⁾.

In the case of the ground state of HH^+ ($n=1$), Sandeman⁽⁹⁾ has found $\omega_e=2321.5 \text{ cm.}^{-1}$, $r_e=1.055 \text{ \AA.U.}$, which gives our constant $k(\text{HH}^+)=\omega_e r_e^3=2660$, which lies near k for the ground state of HH (2520). It will be interesting to pursue the enquiry to ground and excited states of HH^+ , HD^+ , and DD^+ when further experimental results become available.

SUMMARY.

Pursuant to our earlier investigation on non-hydrides, a systematic analysis has been made of the comparative merits of formulæ for calculating equilibrium internuclear distance of hydride di-atoms proposed by Huggins, Allen and Longair, Badger, and ourselves. It is found necessary to take account of electron configuration as well as period and group when assigning suitable constants. With such allowances our formula reproduces the experimental results for hydrides with the least average error (1.4 per cent.). The investigation has been extended to HH, HD, and DD, and to deuterides for which experimental numbers are available, with similar results. A discussion of the influence of isotopes leads to the conclusion that a simple correction in our formula, applied to deuterides, is successful.

References.

- (1) C. H. Douglas Clark and J. L. Stoves, *Phil. Mag.* (7) xxii. p. 1137 (1936); *Trans. Faraday Soc.* xxxiv. p. 1324 (1938).
- (2) C. H. Douglas Clark, *Phil. Mag.* (7) xix. p. 476 (1935).
- (3) R. S. Mulliken, *Chem. Rev.* ix. p. 347 (1931).
- (4) M. L. Huggins, *Jour. Chem. Phys.* iii. p. 473 (1935), iv. p. 308 (1936).
- (5) H. S. Allen and A. K. Longair, *Phil. Mag.* (7) xix. p. 1032 (1935).
- (6) R. M. Badger, *Jour. Chem. Phys.* ii. p. 128 (1934).
- (7) C. H. Douglas Clark, *Phil. Mag.* (7) xviii. p. 459 (1934).
- (8) T. Tanaka and Z. Koana, *Proc. Math. Phys. Soc. Japan*, xvi. p. 365 (1934).
- (9) I. Sandeman, *Proc. Roy. Soc. Edin.* lv. p. 72 (1935).
- (10) C. H. Douglas Clark and J. L. Stoves, '*Nature*,' cxxxvi. p. 682 (1935).
- (11) W. Jevons, '*Report on Band-Spectra of Diatomic Molecules*' (Appendix II.). Camb. Univ. Press (1932).
- (12) H. Sponer, '*Molekülspektren und ihre Anwendung auf chemische Probleme. I. Tabellen.*' Julius Springer, Berlin (1935).

Department of Inorganic Chemistry,
The University of Leeds.

XXXVI. *An Auxiliary Equation for use with the Heaviside Expansion Theorem.*

By W. B. COULTHARD, B.Sc.(Eng.) London, A.M.I.E.E.*

[Received August 29, 1938.]

IN textbooks dealing with the application of Heaviside methods to electrical engineering problems cases of energizing a dead circuit are dealt with fully ; but the reverse problem of breaking a circuit is frequently omitted, or else dealt with by assuming that an E.M.F. equal but opposite to the potential drop across the switch is inserted. As an "ideal" switch is specified, it is frequently difficult to visualize the existence of such a voltage drop. The following development is offered to avoid this difficulty.

(1) *An Auxiliary Equation.*

Let us consider a general equation of the form

$$F(t) = \frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} \dots + a_n y,$$

where the a 's are constant parameters. Adopting the Heaviside notation we get

$$F(t) = p^n y + a_1 p^{n-1} y + \dots + a_n y. \quad (1)$$

Now at $t=0$ let

$$p^{n-1} y = c_1, \quad p^{n-2} y = c_2 \dots \text{etc.},$$

also let

$$p^n y = u(t) \text{ say.}$$

Then by integration we have

$$p^{n-1} y = \int u(t) dt + c_1,$$

$$p^{n-2} y = \int \int u(t) dt^2 + c_1 t + c_2,$$

$$\dots \dots \dots$$

and

$$y = \frac{1}{p^n} u(t) + c_1 \frac{t^{n-1}}{[n-1]} + c_2 \frac{t^{n-2}}{[n-2]} \dots + c_n \quad (2)$$

By substitution of these in equation (1) we get

$$\begin{aligned} F(t) &= u(t) + \frac{a_1}{p} u(t) + \frac{a_2}{p^2} u(t) \dots + \frac{a_n}{p^n} u(t) + a_1 c_1 + a_2 (c_1 t + c_2) \dots \\ &= \left[1 + \frac{a_1}{p} + \frac{a_2}{p^2} \dots \right] u(t) + a_1 c_1 \dots + a_n \left[c_1 \frac{t^{n-1}}{[n-1]} + c_2 \frac{t^{n-2}}{[n-2]} \dots \right]. \quad (3) \end{aligned}$$

* Communicated by the Author.

Now equation (2) may be rewritten as

$$\begin{aligned} u(t) &= p^n \left[y - c_1 \frac{t^{n-1}}{n-1} - c_2 \frac{t^{n-2}}{n-2} \dots - c_n \right] \\ &= p^n [y - f(c)], \\ \text{where} \quad f(c) &= c_1 \frac{t^{n-1}}{n-1} + c_2 \frac{t^{n-2}}{n-2} \dots + c_n \\ &= \frac{c_1}{p^{n-1}} + \frac{c_2}{p^{n-2}} \dots + c_n. \end{aligned}$$

So that we have

$$\begin{aligned} p^n f(c) &= (c_1 p + c_2 p^2 \dots + c_n p^n), \\ a_1 p^{n-1} f(c) &= a_1 c_1 + a_1 c_2 p \dots + a_1 c_n p^{n-1}. \end{aligned}$$

On substituting in equation (3) we have on collecting terms that

$$F(t) = \frac{u(t)}{p^n} [p^n + a_1 p^{n-1} \dots + a_n] - c_1 p - c_2 (p^2 + a_1 p) \dots$$

But $p^{-n}u(t) = y$ by definition.

$$\therefore y = \frac{F(t) + c_1 p + c_2 (p^2 + a_1 p) + \dots}{p^n + a_1 p^{n-1} + a_2 p^{n-2} \dots + a_n} \quad (4)$$

This equation holds equally well for making or breaking a circuit. To show the method we will take the case of opening an (RLC) circuit.

(2) Application.

For the (RLC) circuit the voltage equation is given by

$$Ri + Lpi + \frac{i}{pC} = 0,$$

or

$$p^2 q + \frac{R}{L} p q + \frac{q}{CL} = 0.$$

For boundary conditions we will assume that the charge on the condenser before opening of the switch was steady, i. e., at

$$t=0, \quad q=Q, \quad \frac{dq}{dt}=0.$$

Comparing with equation (1) we have

$$F(t)=0, \quad a_1 = \frac{R}{L}, \quad a_2 = \frac{1}{LC},$$

and from boundary conditions $c_1=0$, $c_2=Q$.

So the form of the auxiliary equation reduces to

$$q = \frac{Q \left(p^2 + \frac{R}{L} p \right)}{p^2 + \frac{R}{L} p + \frac{1}{LC}}.$$

By application of the Expansion Theorem we get

$$q = Q \cdot \sum \frac{\left(p_n + \frac{R}{L}\right)}{\left.\frac{dZ}{dp}\right|_{p=p_n}} \epsilon^{p_n t}.$$

The roots are given by

$$-\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} = -\alpha \pm j\omega_0 \text{ say}$$

for oscillatory conditions.

So we get by the usual methods that

$$q = \frac{Q}{\omega_0 \sqrt{LC}} \epsilon^{-\alpha t} \cos(\omega_0 t - \phi),$$

where

$$\tan \phi = \frac{\alpha}{\omega_0}$$

and

$$i = -\frac{dq}{dt} = \frac{Q \epsilon^{-\alpha t}}{\omega_0 LC} \sin \omega_0 t.$$

Though we have shown the use of the equation when $F(t)=0$, it may be used when $F(t) \neq 0$ and the circuit arrangement is varied by switching in or out parts thereof. It will be noted that when $c_1, c_2 \dots = 0$ then it reduces to the usual equation for an initially dead circuit.

XXXVII. *Approximation Formulæ for a well-known Difference of Products of Two Cylinder Functions.*

By H. BUCHHOLZ *.

[Received September 7, 1938.]

1. *Introduction.*

WHEN calculating the electrical current, the magnetic flux or the energy losses resulting from the inductive effect of a pure sinusoidal alternating current in a circular metal tube with radii R and r cm. ($R > r$), it is eventually necessary to evaluate expressions of the form (1.1 *a*) or (1.1 *b*) in which the cylinder functions are in general of an integral order n .

$$P_{p, n+p}(Rz, rz) = J_n(Rz)Y_{n+p}(rz) - J_{n+p}(rz)Y_n(Rz) \quad (p=0, 1, 2 \dots) \quad (1.1 a)$$

$$P_{n+p, n}(Rz, rz) = J_{n+p}(Rz)Y_n(rz) - J_n(rz)Y_{n+p}(Rz). \quad (1.1 b)$$

In these equations z represents an abbreviation for the product $ki^{1/2} = (\omega\mu\mu_0\sigma i)^{1/2}$ having the dimension $1/\text{cm.}$, and wherein $\omega = 2\pi f$ is the angular frequency of the employed alternating current with the time law $\exp. (-i\omega t)$, μ the unfixed relative permeability, $\mu_0 = 4\pi \cdot 10^{-9} \text{H/cm.}$ the induction constant, and σ the metallic conductivity in S/cm. of the material of which the tube is made. If, for instance, (1.2) represents the component of the exciting vector potential falling along the axis of the tube, then the presence of the tubular conductor will so alter this relation

$$W_z(\rho, \phi) = \sum_{n=1}^{\infty} A_n \left\{ \begin{array}{l} (h/\rho)^n \\ (\rho/h)^n \end{array} \right\} \cdot e^{\pm i n \phi} \quad \begin{array}{l} (\rho > h) \\ (\rho < h) \end{array} \quad (1.2)$$

for the vector potential, according to the position of the excitation, in the space $\rho \geq R > r > h$ or $\rho \leq r < R < h$, that now all coefficients A_n have to be multiplied by factor β_n , this factor

$$\beta_n = -4n/(\pi R r \cdot z^2) \cdot (R/r)^n / P_{n-1, n+1}(Rz, rz) \quad (1.2 a)$$

being given in both cases by (1.2 *a*). Hence factor β_n is a measure for the shielding effect of the space harmonic of the " n th" order, which originates in the thin-walled tube. The more nearly factor β_n approaches unity, the less noticeable this shielding effect becomes.

Direct calculation of the expressions P from its defining equations (1.1 *a*, 1.1 *b*) presents considerable difficulty. Firstly there is a lack

* Communicated by the Author.

of tables giving functions J_n and Y_n with argument $xi^{1/2}$, for higher values of index n . Admittedly for larger values of n , use can already be made of the semi-convergent developments of Debye's (Literature (1)) with some advantage. Since however—and here lies the second and more important reason for the difficulty of reckoning expression P directly,—in most cases in practice, the two radii “R” and “r” differ very little from one another, a fairly accurate determination of the exhaustive function values is already required, in order to obtain to a certain extent an even approximately correct evaluation for the difference of products appearing in (1.1 a) and (1.1 b). For this reason it is attempted in this paper to obtain for the expressions P of (1.1 a) and (1.1 b) other mathematical relationships, permitting a more satisfactory evaluation of them. In so far as these relationships are of only an approximate nature, the approximation covers merely the assumption that for any values of n and z , the difference $R-r$ is always small compared with product $(Rr)^{1/2}$. The formulæ derived present in a most convenient manner the possibility of examining the very various influences of n and z on the behaviour of the function P

2. The Fundamental Integral Representation of the Functions $P_{n,n+p}$.

In developing our desired formulæ, we make use of an integral representation for the product $J_\mu(Rz) J_\nu(rz)$, first announced by A. L. Dixon and W. L. Ferrar (Literature (2)). It consists in the relationship (2.1)

$$\pi J_\mu(Rz) J_\nu(rz) = R^\mu \cdot r^\nu \cdot \int_{-\pi/2}^{+\pi/2} e^{i\theta(\mu-\nu)} \cdot \left[\frac{\lambda_1(\theta)}{\lambda_2(\theta)} \right]^{\mu+\nu} \cdot J_{\mu+\nu}(z\lambda_1(\theta)\lambda_2(\theta)) \cdot d\theta, \quad (2.1)$$

where

$$\lambda_1^2(\theta) = e^{i\theta} + e^{-i\theta} = 2 \cos \theta, \quad \dots \dots \dots (2.1 a)$$

$$\lambda_2^2(\theta) = R^2 \cdot e^{i\theta} + r^2 \cdot e^{-i\theta}. \quad \dots \dots \dots (2.1 b)$$

The proof of (2.1) is very simple. Since it removes the difficulty of using a formula of unknown origin and, besides this, teaches us how to recognize the limits of its application, the derivation of formula (2.1) is included here.

From the definition of the Bessel function we obtain by a simple conversion, the series

$$\begin{aligned} J_\mu(Rz) J_\nu(rz) &= \sum_{r=0}^{\infty} \frac{(-)^r (Rz/2)^{2r+\mu}}{r! \Gamma(\mu+r+1)} \cdot \sum_{s=0}^{\infty} \frac{(-)^s (rz/2)^{2s+\nu}}{s! \Gamma(s+\nu+1)} \quad (m=r+s) \\ &= R^\mu r^\nu \cdot \sum_{m=0}^{\infty} \frac{(-)^m (z/2)^{2m+\mu+\nu}}{m! \Gamma(m+\mu+\nu+1)} \left(\sum_{r=0}^m \frac{m!}{r! (m-r)!} \cdot \frac{\Gamma(m+\mu+\nu+1) R^{2r} r^{2m-2r}}{\Gamma(\mu+r+1) \Gamma(\nu+m-r+1)} \right). \end{aligned}$$

Now for the case $\mu+\nu+m > -1$ (Literature (4))

$$\frac{\Gamma(m+\mu+\nu+1)}{\Gamma(\mu+r+1) \Gamma(\nu+m-r+1)} = \frac{2^{m+\mu+\nu}}{\pi} \cdot \int_{-\pi/2}^{+\pi/2} \cos^{m+\mu+\nu} \theta \cdot e^{i(\mu-\nu+2r-m)\theta} \cdot d\theta.$$

Hence the sum appearing in brackets may be expressed as

$$r^{2m}/\pi \cdot 2^{m+\mu+\nu} \cdot \int_{-\pi/2}^{+\pi/2} \cos^{m+\mu+\nu}\theta \cdot e^{i(\mu-\nu-m)\theta} \left\{ \sum_{r=0}^m \frac{m!}{r!(m-r)!} \left(\frac{R}{r} e^{i\theta} \right)^{2r} \right\} d\theta$$

and here the expression in the curved bracket is identical with

$$(1 + R^2 e^{2i\theta}/r^2)^m = e^{i\theta m/r^2 m} \cdot (R^2 e^{i\theta} + r^2 \cdot e^{-i\theta})^m.$$

A further simplification of the factors in the expression resulting from this operation leads then directly to (2.1), which, providing $\text{Re}(\mu+\nu) > -1$, is conclusive for all real values of R and r and for any complex value of z .

We will now consider (2.1) for the particular case $\mu = -\nu$, from which results the simpler equation (2.2).

$$\pi \cdot J_{-\nu}(Rz) J_{+\nu}(rz) = (r/R)^\nu \int_{-\pi/2}^{+\pi/2} e^{-2i\theta\nu} J_0(z\lambda_1\lambda_2) \cdot d\theta \quad (\nu \neq \pm 0, \pm 1 \dots) \quad (2.2)$$

But from (1.1 a) and (1.1 b) we can now write for the product $\lambda_1^2 \cdot \lambda_2^2$:

$$\lambda_1^2 \cdot \lambda_2^2 = R^2 + r^2 + 2rR \cdot \cos(2\theta - \alpha) \quad \alpha = i \cdot \ln(R/r). \quad (2.1 a)$$

If then in (2) in place of θ we introduce the new integration variable $\phi = -2\theta + \alpha + \pi$, we arrive at

$$2\pi \cdot J_{-\nu}(Rz) J_{\nu}(rz) = e^{-\pi i \nu} \cdot \int_{\alpha}^{2\pi+\alpha} e^{i\nu\phi} \cdot J_0(z\sqrt{R^2+r^2-2Rr \cos \phi}) \cdot d\phi. \quad (2.3)$$

In this equation we exchange R and r and form with (2.3) and the resulting new expression the difference, whence

$$\begin{aligned} e^{-\pi i \nu} \cdot \int_{\alpha}^{2\pi+\alpha} e^{i\nu\phi} \cdot J_0(z\sqrt{R^2+r^2-2Rr \cos \phi}) \cdot d\phi \\ + e^{-\pi i \nu} \cdot \int_{2\pi-\alpha}^{-\alpha} e^{i\nu\phi} \cdot J_0(z\sqrt{R^2+r^2-2Rr \cos \phi}) \cdot d\phi. \quad (2.4) \end{aligned}$$

From fig. 1, however, in consequence of the integral theorem of Cauchy, since the integrand is free from singularities within the rectangle with corners $-\alpha$, $+\alpha$, $2\pi+\alpha$, $2\pi-\alpha$, the sum of the two right-hand integrals may be expressed as the sum of two integrals within limits $-\alpha \dots +\alpha$ and $2\pi-\alpha \dots 2\pi+\alpha$. If in the second of these integrals ϕ is replaced by $\phi+2\pi$, then the right-hand side of (2.4) appears as

$$(e^{2\pi i \nu} - i) \cdot \int_{-\alpha}^{+\alpha} e^{i\nu\phi} \cdot J_0(z\sqrt{R^2+r^2-2Rr \cos \phi}) \cdot dy.$$

Since $\alpha = i \cdot \ln(R/r)$, the variable $i\phi$ is preferable to ϕ in this integral. If the function $J_{-\nu}(Rz)$ is further replaced by functions $Y_{\nu}(Rz)$ and $J_{\nu}(Rz)$ and handled in the same way, then (2.5) appears as the final representation of the difference $P_{n,n}$, wherein now ν , without losing its sense, can

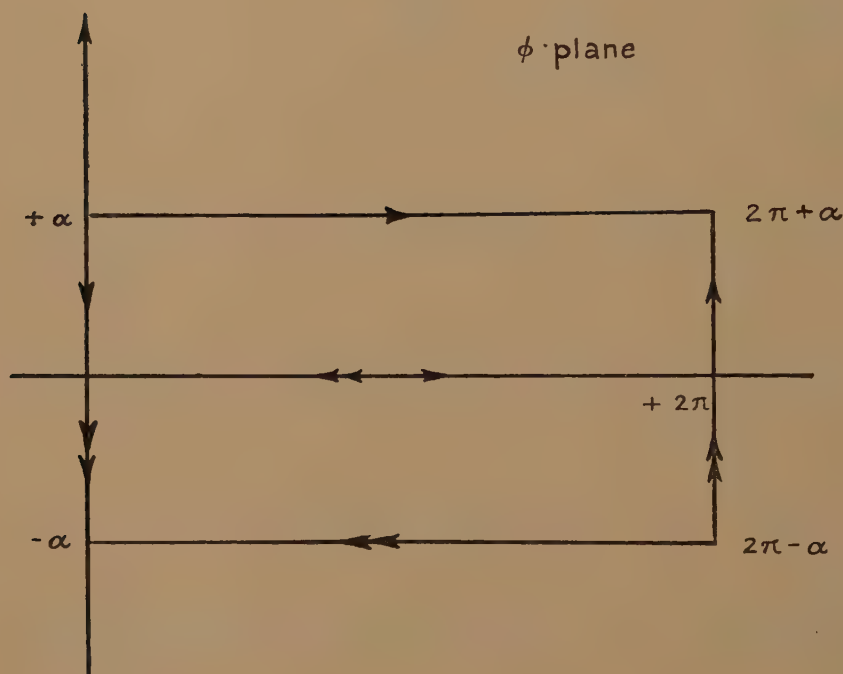
also assume values 0, 1, 2 (whole numbers) as index, which we express by the substitution of n for ν . As it must be, each side of (2.5) equates to zero, when $R=r$, and also each side alters its sign, if R and r are exchanged for each other.

$$P_{n,n}(Rz, rz) = J_n(Rz)Y_n(rz) - J_n(rz)Y_n(Rz)$$

$$= -(1/\pi) \cdot \int_{-\ln(R/r)}^{+\ln(R/r)} e^{n\phi} \cdot J_0(zw_\phi) \cdot d\phi. \quad (2.5)$$

$$w_\phi^2 = R^2 + r^2 - 2Rr \cdot \cosh \phi \quad (2.5a)$$

Fig. 1.



If a partial integration is performed on the integral in (2.5) we then obtain another relationship required for our further progress,

$$P_{n,n}(Rz, rZ) = -\frac{1}{\pi n} \left[\left(\frac{R}{r} \right)^n - \left(\frac{r}{R} \right)^n \right]$$

$$+ \frac{z^2 \cdot Rr}{\pi n} \cdot \int_{-\ln(R/r)}^{+\ln(R/r)} e^{+n\phi} \cdot \frac{J_1(zw_\phi)}{zw_\phi} \cdot \sinh \phi \cdot d\phi. \quad (2.5b)$$

In order to obtain similar representations for the functions $P_{n-1,n}$ and so on, we make use of the well-known recursion formulæ for cylinder

functions. For the function P they take, within the compass here required, the form

$$P_{n-1,n}(Rz, rz) = \frac{n}{Rz} \cdot P_{n,n}(Rz, rz) + \frac{1}{z} \cdot \frac{\partial}{\partial R} P_{n,n}(Rz, rz), \quad (2.6 a)$$

$$P_{n-1,n+1}(Rz, rz) = \begin{cases} -P_{n-1,n-1}(Rz, rz) + \frac{2n}{rz} \cdot P_{n-1,n}(Rz, rz), & (2.6 b_1) \\ -P_{n+1,n+1}(Rz, rz) + \frac{2n}{Rz} \cdot P_{n,n+1}(Rz, rz), & (2.6 b_2) \\ \frac{n}{rz} P_{n-1,n}(Rz, rz) - \frac{1}{z} \cdot \frac{\partial}{\partial r} P_{n-1,n}(Rz, rz), & (2.6 b_3) \end{cases}$$

$$P_{n,n+1}(Rz, rz) = \frac{n}{rz} \cdot P_{n,n}(Rz, rz) - \frac{1}{z} \cdot \frac{\partial}{\partial r} P_{n,n}(Rz, rz). \quad (2.6 c)$$

Adding equations (2.6 b_1) and (2.6 b_2) we obtain for $P_{n-1,n+1}$ the further relationship

$$\begin{aligned} 2P_{n-1,n+1}(Rz, rz) &= -P_{n-1,n-1}(Rz, rz) - P_{n+1,n+1}(Rz, rz) \\ &+ 2 \frac{n}{rz} P_{n-1,n}(Rz, rz) + 2 \frac{n}{Rz} P_{n,n+1}(Rz, rz). \quad (2.6 d) \end{aligned}$$

For the remaining combinations the relation (2.6) which holds for

$$P_{n,n}(Rz, rz) = -P_{n,n}(rz, Rz). \quad (2.6)$$

each pair of indices, is of assistance. Now from equation (2.6) the derivation of $P_{n,n}$ to R or r provides

$$\begin{aligned} \frac{1}{z} \cdot \frac{\partial}{\partial R} P_{n,n}(Rz, rz) &= -\frac{1}{\pi Rz} \left[\left(\frac{R}{r} \right)^n + \left(\frac{r}{R} \right)^n \right] \\ &+ \frac{zR}{\pi} \cdot \int_{-\ln(R/r)}^{+\ln(R/r)} e^{n\phi} \cdot \frac{J_1(zw_\phi)}{zw_\phi} \cdot \left(1 - \frac{r}{R} \cosh \phi \right) d\phi, \quad (2.7 a) \end{aligned}$$

$$\begin{aligned} \frac{1}{z} \cdot \frac{\partial}{\partial r} P_{n,n}(Rz, rz) &= \frac{1}{\pi rz} \left[\left(\frac{R}{r} \right)^n + \left(\frac{r}{R} \right)^n \right] \\ &+ \frac{zr}{\pi} \cdot \int_{-\ln(R/r)}^{+\ln(R/r)} e^{n\phi} \cdot \frac{J_1(zw_\phi)}{zw_\phi} \cdot \left(1 - \frac{R}{r} \cosh \phi \right) d\phi. \quad (2.7 b) \end{aligned}$$

Thus by means of (2.6 a) and (2.5 b) we obtain at once, in addition to (2.5), the relation

$$\begin{aligned} P_{n-1,n}(Rz, rz) &= -\frac{2}{\pi Rz} \cdot \left(\frac{R}{r} \right)^n \\ &+ \frac{rz}{\pi} \cdot \int_{\ln(R/r)}^{+\ln(R/r)} e^{+n\phi} \cdot \frac{J_1(zw_\phi)}{zw_\phi} \cdot \left(\frac{R}{r} - e^{-\phi} \right) \cdot d\phi. \quad (2.8) \end{aligned}$$

In the same manner, after quite elementary working, (2.6 b_3) gives the formula

$$P_{n-1, n+1}(Rz, rz) = -\frac{(R/r)^n}{\pi \cdot Rz \cdot rz} [4n - z^2(R^2 - r^2)] \\ - \frac{r^2 z^2}{\pi} \int_{-\ln(R/r)}^{+\ln(R/r)} e^{(n+1)\phi} \cdot \frac{J_2(zw_\phi)}{z^2 w_\phi^2} \cdot \left(\frac{R}{r} - e^{-\phi}\right)^2 \cdot d\phi. \quad (2.9)$$

In equations (2.5), (2.8), and (2.9) we have collected the required integral representations of the functions $P_{n, n-1}$.

If $R=r$ is inserted in (2.8) or (2.9), then the terms with the integral disappear in these equations, and we obtain

$$P_{n-1, n}(Rz, rz) = J_{n-1}(Rz)Y_n(Rz) - J_n(Rz)Y_{n-1}(Rz) = -2/\pi Rz \quad (2.10 a)$$

$$P_{n-1, n+1}(Rz, Rz) = J_{n-1}(Rz)Y_{n+1}(Rz) - J_{n+1}(Rz)Y_{n-1}(Rz) = -4n/(\pi R^2 z^2) \\ \dots \dots (2.10 b)$$

Both equations (2.10) are well known in the elementary theory of cylinder functions. Besides this, these equations confirm the fact, already well known, that for functions $P_{n, n+p}$, the zero for the z -plane is no longer a branching point, as is yet the case for the functions Y_n . In excess they further provide a full explanation of the functional character of the differences.

3. The Development of the Series after P. Schafheitlin.

Before carrying out further operations on the integrals derived above, in order to obtain the desired approximation formulæ for the function $P_{n, n+p}$, we should show for one of these integrals at least, how to derive the far too little known series, as carried out by P. Schafheitlin. We will take for this purpose the most complicated of the three integrals (2.9) and substitute therein for the Bessel function its equivalent absolute convergent series. The term under the integral may then be written

$$-\frac{r^2 z^2}{4\pi} \cdot \sum_{m=0}^{\infty} \frac{(-)^m (z/2)^{2m}}{m!(m+2)!} \int_{-\ln(R/r)}^{+\ln(R/r)} e^{(n+1)\phi} \\ \cdot (R^2 + r^2 - 2Rr \cdot \cosh \phi)^m \cdot \left(\frac{R}{r} - e^{-\phi}\right)^2 \cdot d\phi. \quad (3.1 a)$$

Then, since $R^2 + r^2 - 2Rr \cdot \cosh \phi = r^2 \cdot (R/r - e^{+\phi})(R/r - e^{-\phi})$ where $e^\phi = t$ the above integral is identical with integral (3.1 b)

$$r^{2m} \cdot \int_{r/R}^{R/r} t^n \cdot (R/r - t)^m \cdot (R/r - 1/t)^{2+m} \cdot dt. \quad (3.1 b)$$

and this, on substituting $1 = r/R + u \cdot (R/r - r/R)$, becomes (3.1 c) :

$$r^{2m} \left(\frac{R^2 - r^2}{r^2}\right)^{2m+3} (r/R)^{n-1} \int_0^1 u^{m+2} (1-u)^m \cdot \left(1 + \frac{R^2 - r^2}{r^2} \cdot u\right)^{n-2-m} \cdot dn \\ \dots \dots (3.1 c)$$

Provided that $(R^2 - r^2)/r^2 < 1$, the integral with variable u in (3.1 c) defines the hypergeometric function

$m!(m+2)!(2m+3)!.F(m+2-n, m+3; 2m+4; 1-R^2/r^2)$,
defined by the equation (3.3),

$$F(\alpha_1, \alpha_2; \rho; X) = 1 + \frac{\alpha_1 \cdot \alpha_2}{\rho} \cdot \frac{X}{1!} + \frac{\alpha_1 \cdot \alpha_1 + 1 \cdot \alpha_2 \cdot \alpha_2 + 1}{\rho \cdot \rho + 1} \cdot \frac{X^2}{2!} + \dots, \quad (3.3)$$

whence appears as the equivalent expression for (3.1 a)

$$-\frac{r^2 z^2}{4\pi} \cdot \left(\frac{r}{R}\right)^{n-1} \cdot \left(\frac{R^2 - r^2}{r^2}\right)^3 \cdot \sum_{m=0}^{\infty} (-)^m \frac{\left(\frac{zr}{2} \cdot \frac{R^2 - r^2}{r^2}\right)^{2m}}{(2m+3)!} \\ F\left(m+2-n, m-3; 2m+4; 1 - \frac{R^2}{r^2}\right). \quad (3.1 d)$$

When $m = -1$, the corresponding term of (3.1 d) has the value $(R/r)^n \cdot (R^2 - r^2)/\pi Rr$. This term, however, already appears in (2.9) before the integral sign. We are therefore enabled to omit this expression, but must then in (3.1 d) begin the summation at $m = -1$. We then substitute for m the new summation variable $p-1$, which then counts from zero, and obtain for $P_{n-1, n+1}$ the final series for $(R^2 - r^2)/r^2 < 1$

$$P_{n-1, n+1}(Rz, rz) = -\frac{4n}{\pi \cdot Rz \cdot rz} \cdot (R/r)^n + (1/\pi) \cdot (r/R)^{n-1} (R^2 - r^2)/r^2 \\ \times \sum_{p=0}^{\infty} (-)^p \cdot \frac{\left(\frac{zr}{2} \cdot \frac{R^2 - r^2}{r^2}\right)^{2p}}{(2p+1)!} \cdot F\left(p+1-n, p+2; 2p+2; 1 - \frac{R^2}{r^2}\right). \quad (3.2 a)$$

If R and r are here transposed, having reference to (2.6) and considering the transformation equation (3.3 a) we derive, in addition to (3.2 a) for the function $P_{n+1, n-1}$, the series

$$F(\alpha_1, \alpha_2; \rho; X) = (1-X)^{-\alpha_1} F\left(\alpha_1, \rho - \alpha_2; \rho; \frac{X}{X-1}\right), \quad (3.3 a)$$

$$P_{n+1, n-1}(Rz, rz) = \frac{4n}{\pi \cdot Rz \cdot rz} \cdot (r/R)^n + 1/\pi \cdot (r/R)^{n-1} \cdot (R^2 - r^2)/R^2 \\ \times \sum_{p=0}^{\infty} (-)^p \cdot \frac{\left(\frac{zr}{2} \cdot \frac{R^2 - r^2}{r^2}\right)^{2p}}{(2p+1)!} \cdot F\left(p+1-n, p; 2p+2; 1 - \frac{R^2}{r^2}\right). \quad (3.2 b)$$

The other integrals (2.5) and (2.8) may be treated in exactly the same manner as integral (2.9). They provide the following series:

$$P_{n, n}(Rz, rz) = -\frac{2}{\pi rz} \cdot (r/R)^n \\ \times \sum_{p=0}^{\infty} (-)^p \frac{\left(\frac{zr}{2} \cdot \frac{R^2 - r^2}{r^2}\right)^{2p+1}}{(2p+1)!} \cdot F\left(p+1-n, p+1; 2p+2; 1 - \frac{R^2}{r^2}\right), \quad (4)$$

$$P_{n-1, n}(Rz, rz) = -\frac{2}{\pi rz} (r/R)^{n-1} \times \sum_{p=0}^{\infty} (-)^p \frac{\left(\frac{zr}{2} \cdot \frac{R^2-r^2}{r^2}\right)^{2p}}{2p!} \cdot F\left(p+1-n, p+1; 2p+1; 1-\frac{R^2}{r^2}\right), \quad (5a)$$

$$P_{n, n-1}(Rz, rz) = \frac{2}{\pi rz} (r/R)^n \times \sum_{p=0}^{\infty} (-)^p \frac{\left(\frac{zr}{2} \cdot \frac{R^2-r^2}{r^2}\right)^{2p}}{2p!} \cdot F\left(p+1-n, p; 2p+1; 1-\frac{R^2}{r^2}\right). \quad (5b)$$

A few of these series, which without exception hold only when $(R^2-r^2)/r^2 < 1$, were at first obtained by P. Schafheitlin (Literature (3)) in a quite different way. They lend themselves very well to the calculation of the function P , particularly in the present case of practically equal values of r and R , since on this assumption they converge very rapidly.

4. Transformation of the Integrals so far obtained and their Approximation Formulæ.

The integrals appearing in (2.5), (2.8), and (2.9) are always of the same type as (4.1). Now from the explanatory equation (2.5 a) for

$$X_n^{(p)}(z; R, r) = \int_{-\ln(R/r)}^{+\ln(R/r)} e^{+n\phi} \cdot \frac{J_p(zw_\phi)}{(zw_\phi)^p} \cdot d\phi \quad (p=0, 1, 2) \quad (4.1)$$

w_ϕ , this quantity may also be written in the form :

$$w_\phi^2 = R^2 + r^2 - 2Rr \cdot \cosh \phi = (R-r)^2 \cdot \left[1 - \frac{4Rr}{(R-r)^2} \cdot \sinh^2 \frac{\phi}{2}\right]. \quad (4.2)$$

We then insert the two dimensionless auxiliary quantities ζ and δ according (4.3 a, b), and, further, in (4.1) we make use of the new integration variable η defined by (4.3)

$$\delta = (R-r)/\sqrt{Rr}, \quad (4.3a)$$

$$\zeta = z \cdot \sqrt{Rr} \quad (4.3b)$$

$$\sinh \frac{\phi}{2} = (\delta/2) \cdot \sin \eta. \quad (4.3)$$

These substitutions put (4.1) into the form

$$X_n^{(p)}(z; R, r) = \delta \cdot \int_{-\pi/2}^{+\pi/2} e^{2n \cdot \text{arc sinh}(\delta/2 \cdot \sin \eta)} \cdot \frac{J_p(\zeta \delta \cdot \cos \eta)}{(\zeta \delta \cdot \cos \eta)^p} \cdot \frac{\cos \eta \cdot d\eta}{\sqrt{1 + (\delta^2/4) \cdot \sin^2 \eta}}. \quad (4.4)$$

Now in most cases occurring in practice r is nearly equal to R , and therefore δ is so small that the denominator in (4.4), $1 + (\delta^2/4) \cdot \sin^2 \eta$, may

without appreciable error be considered as equal to unity. From (4.5) with the same degree of accuracy, it then becomes permissible to substitute

$$\operatorname{arc} \sinh x = x \cdot F\left(\frac{1}{2}; \frac{1}{2}; \frac{3}{2}; -x^2\right) = x - \frac{x^3}{6} + \frac{3}{40}x^5 \dots \quad (4.5)$$

$\delta/2 \cdot \sin \eta$ for the exponent of the e -function $\operatorname{arc} \sinh (\delta/2 \cdot \sin \eta)$. On the other hand, the fraction $J_p(\zeta \delta \cos \eta)/(\zeta \delta \cdot \cos \eta)^p$, which for $\delta \rightarrow 0$ approaches the value $1/(p! 2^p)$, remains practically unaffected by the size of δ , if small. Consequently, we have for small δ independent of the values of n and z the approximation formula

$$X_n^{(p)}(z, R, r) = \delta \cdot \int_{-\pi/2}^{+\pi/2} e^{n\delta \cdot \sin \eta} \cdot \frac{J_p(\zeta \delta \cdot \cos \eta)}{(\zeta \delta \cdot \cos \eta)^p} \cdot \cos \eta \cdot d\eta + O(\delta^3). \quad (4.4 a)$$

Thus all integrals are small of the same order as δ , and the first term neglected in the transformation (4.4) into (4.4 a) is of the order δ^3 .

The integral of (4.4 a) may be given in a finite form for all values of p . For the practical utilization of the integrals, however, which we have mainly in view, only the particular three corresponding to the values 0, 1 and 2 of p are of interest. It is, however, quite unnecessary to evaluate each of the three separately, since a simple relationship exists between two successive expressions. From the integral representations of (4.4 a), the following two relationships may be directly stated:—

$$X_n^{(0)}(z; R, r) = 2\delta \cdot \frac{\sinh(n\delta)}{n\delta} - \frac{(\zeta\delta)^2}{n\delta} \cdot \frac{\partial(X_n^{(1)})}{\partial(n\delta)}, \quad \dots \quad (4.6 a)$$

$$X_n^{(1)}(z; R, r) = \delta \cdot \frac{\sinh(n\delta)}{n\delta} - \frac{(\zeta\delta)^2}{n\delta} \cdot \frac{\partial(X_n^{(2)})}{\partial(n\delta)}. \quad \dots \quad (4.6 b)$$

In both these equations the differentiation refers only to the product $n\delta$ as a whole. On account of this relationship we need only to consider the calculation of integral (4.4 a) for $p=2$. If, for the sake of brevity, we substitute a for $n\delta$ and b for $\zeta\delta$, and further replace the exponential function and the Bessel function under the integrand by its equivalent absolute convergent series, we obtain, since the \sin is an odd function, in place of (4.4 a) the expression

$$X_n^{(2)}(z; R, r) = 2\delta/b^2 \cdot \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} \frac{a^{2n}}{2n!} \cdot \frac{(-)^p \cdot (b/2)^{2+2p}}{p!(p+2)!} \times \int_0^{\pi/2} \sin^{2n}\phi \cdot \cos^{2p+1}\phi \cdot d\phi \quad \dots \quad (4.7 a)$$

But according to Nielsen, p. 133 (Literature (4)), and from the duplication formula for the Γ -function

$$\int_0^{\pi/2} \sin^{2n}\phi \cdot \cos^{2p+1}\phi \cdot dy = \frac{1}{2} \cdot \frac{\Gamma(n+\frac{1}{2}) \cdot p!}{\Gamma(p+n+\frac{3}{2})}, \quad \dots \quad (4.7 a)$$

$$2n! = \Gamma(2n+1) = \pi^{-\frac{1}{2}} \cdot 2^{2n} \cdot n! \cdot \Gamma(n+\frac{1}{2}), \quad \dots \quad (4.7 b)$$

whence also

$$\begin{aligned} X_n^{(2)}(z; R, r) &= \delta/b^2 \cdot \sqrt{\pi} \cdot \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} \frac{(a/2)^{2n}}{n!} \cdot \frac{(-)^p (b/2)^{2+2p}}{(p+2)! \Gamma(p+n+\frac{3}{2})} \\ &= -\sqrt{\pi} \cdot \delta/b^2 \cdot \sum_{n=0}^{\infty} \sum_{p=1}^{\infty} \frac{(a/2)^{2n} \cdot (ib/2)^{2p}}{n! (p+1)! \Gamma(p+n+\frac{1}{2})}. \quad (4.7 b) \end{aligned}$$

For $p+n=r$, after a separation of a simple series and an inversion of the double series, the right-hand side of (4.7 b) becomes

$$\begin{aligned} & -\sqrt{\pi} \cdot \delta/b^2 \cdot \sum_{n=0}^{\infty} \sum_{r=n+1}^{\infty} \frac{(a/ib)^{2n} \cdot (ib/2)^{2r}}{n! \Gamma(r+\frac{1}{2})(r+1-n)!} \quad (p+n=r) \\ &= \sqrt{\pi} \cdot \delta/b^2 \cdot \sum_{n=0}^{\infty} \frac{(a/2)^{2n}}{n! \Gamma(n+\frac{1}{2})} - \sqrt{\pi} \cdot \delta/b^2 \cdot \sum_{n=0}^{\infty} \sum_{r=n}^{\infty} \frac{(a/ib)^{2n} \cdot (ib/2)^{2n}}{n! (r+1-n)! \Gamma(r+\frac{1}{2})} \\ &= \delta/b^2 \cdot \cosh a - \sqrt{\pi} \cdot \delta/b^2 \cdot \sum_{r=0}^{\infty} \sum_{n=0}^r \frac{(a/ib)^{2n} \cdot (ib/2)^{2n}}{n! (r+1-n)! \Gamma(r+\frac{1}{2})} \\ &= \delta/b^2 \cdot \cosh a - \sqrt{\pi} \cdot \delta/b^2 \cdot \sum_{r=0}^{\infty} \frac{(ib/2)^r}{\Gamma(r+\frac{1}{2}) \cdot (r+1)!} \\ & \quad \times \left(\sum_{n=0}^r (-)^n \frac{(r+1)!}{n! (r+1-n)!} \left(\frac{a^2}{b^2} \right)^n \right). \end{aligned}$$

And now after addition and subtraction of the term $n=r+1$,

$$\sum_{n=0}^r (-)^n \frac{(r+1)!}{n! (r+1-n)!} \left(\frac{a^2}{b^2} \right)^n = (-)^r \cdot \left(\frac{a}{b} \right)^{2r+2} + \left(1 - \frac{a^2}{b^2} \right)^{r+1}.$$

Consequently,

$$\begin{aligned} X_n^{(2)}(z; R, r) &= \delta/b^2 \cdot \cosh a - \sqrt{\pi} \cdot \delta/b^2 \cdot a^2/b^2 \cdot \sum_{r=0}^{\infty} \frac{(a^2/4)^r}{\Gamma(r+\frac{1}{2}) \Gamma(r+2)} \\ & \quad - \sqrt{\pi} \cdot \delta/b^2 \cdot \left(1 - \frac{a^2}{b^2} \right) \cdot \sum_{r=0}^{\infty} ((a^2-b^2)/4)^r / \Gamma(r+\frac{1}{2}) \Gamma(r+2) \quad (4.7 c) \end{aligned}$$

From the definition of a Bessel function, however, we have the relationship

$$\begin{aligned} \sum_{r=0}^{\infty} \frac{X^r}{\Gamma(r+\frac{1}{2}) \Gamma(r+2)} &= \frac{1}{2X\sqrt{\pi}} + \frac{(i\sqrt{X})^{3/2}}{X} \cdot J_{-3/2}(2i\sqrt{X}) \\ &= \frac{1}{2X\sqrt{\pi}} [1 + 2\sqrt{X} \cdot \sinh(2\sqrt{X}) - \cosh(2\sqrt{X})]. \end{aligned}$$

Hence, with the old symbols, the following final representation is obtained:

$$\begin{aligned} X_n^{(2)}(z; R, r) &\approx \frac{2\delta}{\zeta^4 \delta^4} \cdot [\tfrac{1}{2} \zeta^2 \delta^2 \cdot \cosh n\delta + \cosh n\delta - n\delta \cdot \sinh n\delta \\ & \quad - \cosh(\delta\sqrt{n^2 - \zeta^2}) + \delta\sqrt{n^2 - \zeta^2} \cdot \sinh(\delta\sqrt{n^2 - \zeta^2})] + O(\delta^3). \quad (4.8 a) \end{aligned}$$

In this equation as for integral (4.4 a), $X_n^{(2)}$ itself approaches zero as $\delta \rightarrow 0$, since the square bracket disappears as δ^4 .

In order to obtain the corresponding equations for the integrals $X_n^{(1)}$ and $X_n^{(0)}$, we need only to apply (4.6 a, b). This gives directly

$$X^{(1)}(z; R, r) = \frac{2\delta}{\zeta^2 \delta^2} \cdot [\cosh n\delta - \cosh(\delta\sqrt{n^2 - \zeta^2})] + O(\delta^3), \quad (4.8 b)$$

$$X_n^{(0)}(z; R, r) = 2\delta \cdot \sinh(\delta\sqrt{n^2 - \zeta^2})/(\delta\sqrt{n^2 - \zeta^2}) + O(\delta^3), \quad (4.8 c)$$

and thus the approximate calculation of the integral (4.1) for the values $p=0, 1$ and 2 mainly occurring in practice, is finally accomplished.

5. Approximation Formulæ for the Functions $P_{n, n+p}$

From the explanatory equation (4.1) for the integrals $X_n^{(p)}$, the initially defined functions $P_{n, n}$ and so on, may be represented in the following manner, according to (2.5), (2.8), and (2.9) :

$$P_{n, n}(Rz, rz) = -(1/\pi) \cdot X_n^{(0)}(z; R, r), \quad . \quad . \quad (5.1 a)$$

$$P_{n-1, n}(Rz, rz) = -\frac{2}{\pi Rz} (R/r)^n + \frac{Rz}{\pi} \cdot X_n^{(1)}(z; R, r) - \frac{rz}{\pi} \cdot X_{n-1}^{(1)}(z; R, r), \quad . \quad . \quad (5.1 b)$$

$$P_{n-1, n+1}(Rz, rz) = -\frac{(R/r)^n}{\pi \cdot Rz \cdot rz} \cdot [4n - z^2(R^2 - r^2)] - \frac{R^2 z^2}{\pi} \cdot X_{n+1}^{(2)}(z; R, r) \\ - \frac{r^2 z^2}{\pi} X_{n-1}^{(2)}(z; R, r) + 2 \cdot \frac{Rr \cdot z^2}{\pi} X_n^{(2)}(z; R, r), \quad (5.1 c_1)$$

and from (2.6 d) the function $P_{n-1, n+1}$ may also be expressed by another relationship,

$$P_{n-1, n+1}(Rz, rz) = (1/2\pi) \cdot [X_{n-1}^{(0)}(z; R, r) + X_{n+1}^{(0)}(z; R, r)] - \frac{4n \cdot (R/r)^n}{\pi \cdot Rz \cdot rz} \\ + \frac{n}{\pi} \cdot \left[\frac{R^2 - r^2}{r^2} \cdot X_n^{(1)}(z; R, r) - X_{n-1}^{(1)}(z; R, r) + X_{n+1}^{(1)}(z; R, r) \right]. \quad (5.1 c_2)$$

In these equations we make use of the above formulæ (4.8 a-c) and for the quantity $P_{n, n+p}$ itself, we then obtain the following approximation formulæ :

$$P_{n, n}(Rz, rz) = -\frac{2\delta}{\pi} \cdot \frac{\sinh(\delta\sqrt{n^2 - \zeta^2})}{\delta\sqrt{n^2 - \zeta^2}} + O(\delta^3), \quad . \quad . \quad (5.2)$$

$$P_{n-1, n}(Rz, rz) = +\frac{2}{\pi \zeta \delta} \cdot \left[\sqrt{\frac{r}{R}} \cdot \cosh(\delta\sqrt{(n-1)^2 - \zeta^2}) \right. \\ \left. - \sqrt{\frac{R}{r}} \cdot \cosh(\delta\sqrt{n^2 - \zeta^2}) \right] + O(\delta^3), \quad (5.3)$$

$$P_{n-1, n+1}(Rz, rz) = \frac{\delta}{\pi} \cdot \left[\frac{\sinh(\delta\sqrt{(n+1)^2 - \zeta^2})}{\delta\sqrt{(n+1)^2 - \zeta^2}} + \frac{\sinh(\delta\sqrt{(n-1)^2 - \zeta^2})}{\delta\sqrt{(n-1)^2 - \zeta^2}} \right] \\ - \frac{4n}{\pi\zeta^2} \left[\cosh(\delta\sqrt{n^2 - \zeta^2}) + \frac{1}{2\delta} (\cosh(\delta\sqrt{(n+1)^2 - \zeta^2}) \right. \\ \left. - \cosh(\delta\sqrt{(n-1)^2 - \zeta^2})) \right] + O(\delta^3). \quad (5.4)$$

In these equations use has already been made of the fact that, owing to the smallness of δ and to the limited degree of accuracy, $\sinh \delta \approx \delta$ and $\cosh \delta \approx 1$ may be substituted. Moreover, from (4.3 a) we have the equation

$$(R/r)^{\frac{1}{2}} = \delta/2 + \sqrt{1 + \delta^2/4} = e^{\text{arc sinh } (\delta/2)} \approx e^{\delta/2}. \quad (5.4 a)$$

Consequently, $(R/r)^p$ may be equated to $e^{p\delta}$ for any real value of r , and this substitution includes even the term δ^2 .

With reference to the approximation (5.4), the shielding coefficient defined by (1.2 a) in particular may be written

$$1/\beta_n = e^{-n\delta} \cdot \left\{ \cosh(\delta\sqrt{n^2 - \zeta^2}) \right. \\ \left. + \frac{1}{2\delta} (\cosh(\delta\sqrt{(n+1)^2 - \zeta^2}) - \cosh(\delta\sqrt{(n-1)^2 - \zeta^2})) \right\} \\ - \zeta^2\delta/4n \cdot e^{-n\delta} \cdot \left[\frac{\sinh(\delta\sqrt{(n+1)^2 - \zeta^2})}{\delta\sqrt{(n+1)^2 - \zeta^2}} + \frac{\sinh(\delta\sqrt{(n-1)^2 - \zeta^2})}{\delta\sqrt{(n-1)^2 - \zeta^2}} \right], \quad (5.5)$$

when $\delta \rightarrow 0$ β_n has, as might be expected, the limiting value 1.

When reckoning numerically the functions $P_{n, n+p}$ or factor β_n with the assistance of the deduced approximations, the process is aggravated by the fact that, in the equations which we have in mind, the quantity $\zeta^2 = Rr \cdot z^2 = Rr \cdot k^2 i$ is not real, but purely imaginary. The expressions of the form $\delta\sqrt{n^2 - \zeta^2}$ must consequently be separated into their real and imaginary components.

Thus if

$$\sqrt{n^2 - \zeta^2} = \sqrt{n^2 - Rr \cdot k^2 \cdot i} = \alpha_1 - \alpha_2 \cdot i, \quad (5.6)$$

we obtain for the components α_1 and α_2 the two other equations

$$\alpha_1 = (2^{-1/2}) \cdot \{ +n^2 + \sqrt{n^4 + (rRk^2)^2} \}^{1/2}, \quad (5.6 a)$$

$$\alpha_2 = (2^{-1/2}) \cdot \{ -n^2 + \sqrt{n^4 + (rRk^2)^2} \}^{1/2}, \quad (5.6 b)$$

and then

$$\sinh(\delta\sqrt{n^2 - \zeta^2}) = \sinh(\alpha_1\delta) \cdot \cos(\alpha_2\delta) - i \cdot \cosh(\alpha_1\delta) \cdot \sin(\alpha_2\delta), \quad (7a)$$

$$\cosh(\delta\sqrt{n^2 - \zeta^2}) = \cosh(\alpha_1\delta) \cdot \cos(\alpha_2\delta) - i \cdot \sinh(\alpha_1\delta) \cdot \sin(\alpha_2\delta). \quad (7 b)$$

By means of (5.5) the behaviour of the quantity β_n can be easily examined in those cases where n is comparatively large compared with ζ or *vice versa*. From a formula given by Lommel (Literature (5)), for $\alpha^2 < 1$

$$\frac{J_\nu(z\sqrt{1-\alpha^2})}{(z\sqrt{1-\alpha^2})^\nu} = \sum_{m=0}^{\infty} \frac{(\frac{1}{2}\alpha^2 z)^m}{m!} \cdot \frac{J_{\nu+m}(z)}{z^\nu} \quad (5.8)$$

In the particular cases where $\nu = -1/2$ and $\nu = +1/2$, and for sufficiently small values of α^2 with $z = i \cdot u$, we obtain the approximations

$$\cosh(u\sqrt{1-\alpha^2}) \approx \cosh u - \frac{\alpha^2}{2} \cdot u \cdot \sinh u, \quad (5.8a)$$

$$\frac{\sinh(u\sqrt{1-\alpha^2})}{\sqrt{1-\alpha^2}} \approx \frac{\sinh u}{u} + \frac{\alpha^2}{2} \cdot \left(\frac{\sinh u}{u} - \cosh u \right). \quad (5.8b)$$

If, in the first case, n is considerably larger than $|\zeta|$, $n\sqrt{1-(\zeta/n)^2}$ can be written for $\sqrt{n^2-\zeta^2}$, and here the quotient ζ/n fulfils the condition of α .

From (4.8 a), however,

$$\begin{aligned} & \frac{1}{2\delta} [\cosh(\delta\sqrt{(n+1)^2-\zeta^2}) - \cosh(\delta\sqrt{(n-1)^2-\zeta^2})] \\ & \approx \sinh n\delta + \frac{\zeta^2/2}{n^2-1} (\sinh n\delta - n\delta \cdot \cosh n\delta) \\ & \cosh(\delta\sqrt{n^2-\zeta^2}) \approx \cosh n\delta - \frac{1}{2} \cdot (\zeta/n)^2 \cdot n\delta \cdot \sinh(n\delta), \end{aligned}$$

and from (5.5), still considering the relationship $r/R \approx \exp(-\delta)$ for those small or large values of n for which $|\zeta|/n$ is $\ll 1$, we derive

$$1/\beta_n \approx 1 - \frac{1}{2} \cdot (\zeta/n) \cdot \zeta\delta. \quad (5.9)$$

The factor β_n , determining the shielding effect, therefore approaches unity more and more nearly as n increases. For those space harmonics of the exciting field wherein n considerably exceeds the absolute value of ζ , the shielding effect arising from the thin-walled metal tube is consequently always inconsiderable.

If, in the second case, however, the generally more apposite condition is made that $n/|\zeta| \ll 1$, then, with the same degree of accuracy for the coefficient of shielding, we obtain

$$\begin{aligned} 1/\beta & \sim -\zeta/2n \cdot \sin \delta\zeta \cdot e^{-n\delta} \cdot \left[\frac{1-\frac{3}{2}}{\zeta^2} \cdot \frac{n^2}{\zeta^2} \left(1 + \frac{2}{3}n\delta - \frac{1}{3n^2} \right) \right] \\ & + \cos \delta\zeta \cdot e^{-n\delta} \cdot \left(1 + \delta \cdot \frac{n^2+1}{4n} \right). \quad (5.10) \end{aligned}$$

Therefore, for large values of ζ , β_n is very small.

Literature.

- (1) E. Jahnke and F. Emde, 'Funktionentafeln mit Formeln und Kurven,' 2. Auflage, Leipzig, p. 203 (1933).
- (2) A. L. Dixon and W. L. Ferrar, "Integrals for the Product of two Bessel Functions," 'Quarterly Journal of Mathematics,' Oxford, ser. iv. pp. 193-201 (1933).
- (3) P. Schafheitlin, 'Die Theorie des Besselschen Funktionen,' Leipzig, pp. 85 & 86 (1908).
- (4) N. Nielsen, 'Handbuch der Theorie der Gammafunktion,' Leipzig, pp. 158 & 133 (1906).
- (5) G. N. Watson, 'Theory of Besselfunctions,' Cambridge, p. 141 (1922).

Investigation Department of the
AEG-Central Laboratory for Electrical
Communication.

XXXVIII. *Diffraction and Refraction of a Horizontally Polarized
Electromagnetic Wave over a Spherical Earth.*

By MARION C. GRAY *.

[Received November 4, 1938.]

ABSTRACT.

Formulas are derived for the electromagnetic field at a point on or above the surface of a spherical earth due to the presence of a vertical magnetic dipole. It is shown that the resultant field resembles that due to a vertical electric dipole above a spherical earth of low conductivity, and that in the magnetic case the values of the earth constants are of much less importance than in the electric. Curves are included showing the variation of the field with distance and with height.

1. *Introduction.*

IN recent papers van der Pol and Bremmer⁽¹⁾ and Wwedensky⁽²⁾ have given exhaustive treatments of the electromagnetic field generated by a vertical electric dipole situated above a homogeneous earth, assumed to be a sphere of finite radius. Their work is, of course, based on a fundamental paper by Watson⁽³⁾, who showed how a theoretical solution in the form of a very slowly converging series of harmonics could be made of practical utility by transforming it to a much more rapidly converging asymptotic series.

The corresponding problem for a horizontal electric dipole involves a further difficulty in that symmetry about the axis of the dipole can no longer be assumed. A rigorous treatment of the horizontal dipole would thus involve consideration of all three spherical components of both the electric and the magnetic fields.

The present paper is an attempt to obviate this difficulty, and at the same time to obtain an approximate solution of the problem of horizontal polarization, by considering the electromagnetic field generated above the surface of a spherical earth by a small horizontal loop carrying an electric current. It is well known that such a loop is essentially equivalent to a fictitious magnetic dipole perpendicular to the plane of the loop. By considering this vertical magnetic dipole symmetry about the axis of the transmitter is again obtained.

* Communicated by the Author.

2. Theoretical Solution.

Mathematically, then, the problem to be considered reduces to the following (see fig. 1).

Given a vertical magnetic dipole of moment Ml at a point O ($b, 0, 0$), above a sphere of radius a , to determine the electromagnetic field at an arbitrary point P (r, θ, ϕ), on or above the surface of the sphere and at a distance $D = (r^2 + b^2 - 2br \cos \theta)^{1/2}$ from O .

Watson's method can be applied directly to the problem thus formulated, so that a brief outline of the solution will be sufficient.

Fig. 1.

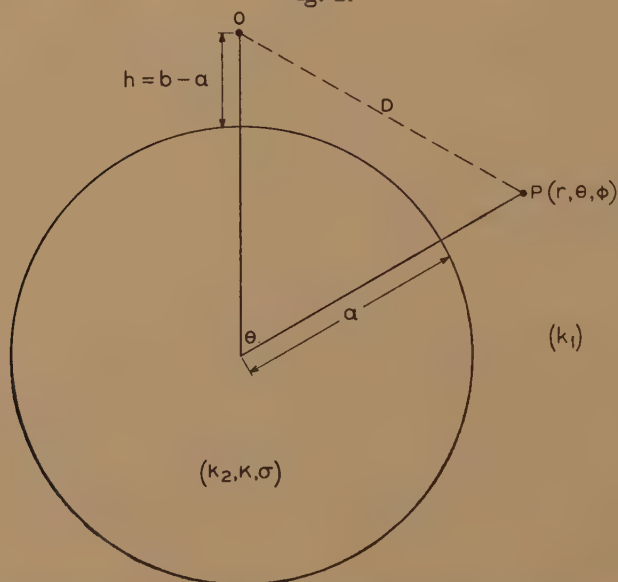


Diagram of co-ordinate system, with transmitter at O , receiver at P .

The primary field of the magnetic dipole can, as usual, be expressed in terms of a Hertzian vector *

$$\Pi_{\text{prim}} = \Pi_r = \frac{Ml}{4\pi i \omega \mu} \frac{e^{-ik_1 D}}{D}.$$

* In conformity with the usual engineering practice a time factor $e^{i\omega t}$ is to be understood throughout. The formulas obtained are thus in line with the original Watson formulas rather than with those of references ⁽¹⁾ and ⁽²⁾, which are based on the conjugate complex time factor $e^{-i\omega t}$. Further, the units used are the m.k.s. practical units, and all formulas are rationalized. Then in any medium the propagation constant is $ik = [+i\omega\mu(\sigma + i\omega\epsilon)]^{1/2}$, so that, in air, $k_1 = 2\pi/\lambda$, while in the earth $k_2 = k_1 (K - 60i\sigma\lambda)^{1/2}$, where K is the relative dielectric constant, σ the conductivity (in mhos per meter), and the permeability μ is assumed to be the same in both media.

For convenience, this is written in the form

$$\Pi_{\text{prim}} = A \frac{e^{-ik_1 D}}{ik_1 D}, \quad \dots \quad (1)$$

where A is proportional to the magnetic moment of the dipole, its value being

$$A = \frac{MI}{4\pi Z} = ik_1 \frac{IS}{4\pi},$$

Z being the characteristic or intrinsic impedance of free space*, and I the electric current in the initial primary loop of area S .

In terms of Π the electric and magnetic fields are:—

$$\left. \begin{aligned} E_r = E_\theta = H_\phi = 0, \quad E_\phi = -i\omega\mu \frac{\partial \Pi}{\partial \theta}, \\ H_\theta = -\frac{1}{r} \frac{\partial^2 (r\Pi)}{\partial r \partial \theta}, \quad H_r = -\left(k^2 + \frac{\partial^2}{\partial r^2}\right)(r\Pi). \end{aligned} \right\} \dots \quad (3)$$

The primary field (1) is now expanded in a series of harmonics,

$$\Pi_{\text{prim}} = A \sum_{m=0}^{\infty} (2m+1) \psi_m(k_1 r) \zeta_m(k_1 b) P_m(\cos \theta), \quad a \leq r \leq b, \quad \dots \quad (4)$$

where

$$\psi_m(x) = \left(\frac{\pi}{2x}\right)^{1/2} J_{m+1/2}(x), \quad \zeta_m(x) = \left(\frac{\pi}{2x}\right)^{1/2} H_{m+1/2}^{(2)}(x), \quad \dots \quad (5)$$

and secondary fields in earth and air are assumed of the same form, but with arbitrary constants to be determined by the boundary conditions. These are the continuity of the tangential components of E and H ,

equivalent to the continuity of Π and $\frac{\partial}{\partial r}(r\Pi)$, at the surface of the earth.

When these conditions are satisfied the total field in the air is found to be:—

$$\Pi = \Pi_{\text{prim}} + \Pi_{\text{sec}} = A \sum_{m=0}^{\infty} (2m+1) \zeta_m(k_1 b) \frac{F_m}{D_m} P_m(\cos \theta), \quad \dots \quad (6)$$

where

$$F_m = k_1 a \psi_m(k_2 a) [\psi_m(k_1 r) \zeta'_m(k_1 a) - \zeta_m(k_1 r) \psi'_m(k_1 a)] \\ + k_2 a \psi'_m(k_2 a) [\psi_m(k_1 a) \zeta'_m(k_1 r) - \psi_m(k_1 r) \zeta'_m(k_1 a)],$$

$$D_m = k_1 a \psi_m(k_2 a) \zeta'_m(k_1 a) - k_2 a \psi'_m(k_2 a) \zeta_m(k_1 a).$$

* The term "characteristic impedance" was adapted from the terminology of network theory by G. A. Campbell ('Physics', v. 3, pp. 230–239, Nov. 1932); in a more recent discussion of impedance concepts S. A. Schelkunoff (Bell System Technical J. v. 17, pp. 17–48, Jan. 1938) has suggested that the term "intrinsic impedance" should be used for the free space value of Z . Its numerical value in m.k.s. units is $(\mu/\epsilon)^{1/2}$, or very nearly 120π ohms.

3. Practical Solution for Zero Heights.

Equation (6) constitutes a complete theoretical solution of the magnetic dipole problem, but in developing the practical solution the case of both transmitter and receiver on the surface of the earth will be considered first. Substituting $r=b=a$ the formula for Π becomes

$$\Pi = -i \frac{A}{z_1^2} \sum_{m=0}^{\infty} (2m+1) \frac{1}{N_m} P_m(\cos \theta), \dots \dots \dots (7)$$

where

$$N_m = \frac{\zeta'_m(z_1)}{\zeta_m(z_1)} - \frac{k_2 \psi'_m(z_2)}{k_1 \psi_m(z_2)},$$

and

$$z_1 = k_1 \bar{a}, \quad z_2 = k_2 \bar{a}. \quad \dots \dots \dots (8)$$

The quantity \bar{a} is introduced instead of a to denote the effective radius of the earth when refraction is taken into account. It has been shown by Schelleng, Burrows and Ferrell⁽⁴⁾ and by Eckersley and Millington⁽⁵⁾ that the refractive effect of the earth's atmosphere can be compensated for by replacing the actual radius a by an effective value \bar{a} , which is a constant multiple of a depending on atmospheric conditions. The explicit dependence will be indicated by writing $\bar{a} = \rho a$ where an average value of ρ is about 4/3, though it may vary between limits of 1 and 3. This convention will be retained throughout, so that the effect of refraction is included in all formulas.

It may be noted also that if the earth were assumed to be of perfect conductivity the field at the surface would be zero. Hence a finite value of the conductivity must be retained, at least in the present case of zero heights.

Watson's method of transforming the series (7) to a contour integral, then deforming the contour to express the integral as the sum of a series of residues at new poles of the integrand is now adopted, and the resulting series is found to be

$$\Pi = i 2 \pi \frac{A}{z_1^2} \sum_s \frac{\nu_s}{\cos \nu_s \pi} \frac{P_{\nu_s-1/2}(-\cos \theta)}{\left[\frac{\partial N}{\partial \nu} \right]_{\nu=\nu_s}}, \dots \dots \dots (9)$$

where ν_s is the s^{th} root of the "residue equation"

$$N \equiv \frac{\zeta'_{\nu-1/2}(z_1)}{\zeta_{\nu-1/2}(z_1)} - \frac{k_2 \psi'_{\nu-1/2}(z_2)}{k_1 \psi_{\nu-1/2}(z_2)} = 0. \quad \dots \dots \dots (10)$$

The important values of ν are those in the neighbourhood of $\nu = z_1$, so that in the series (9) the asymptotic expansion

$$\frac{P_{\nu_s-1/2}(-\cos \theta)}{\cos \nu_s \pi} = \left(\frac{2}{\pi \nu_s \sin \theta} \right)^{1/2} e^{-i \nu_s \theta - i \frac{\pi}{4}}$$

will be valid even for small angles θ as long as $v\theta$ is not small, *i. e.*, at distances which are large compared with the wave-length. Introducing this expansion, and also the additional parameter τ_s (already used by van der Pol and Bremmer) defined by

$$\nu_s = z_1 + \tau_s z_1^{1/3}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

into the series (9), it reduces to

$$P = \frac{2A}{z_1^2} \left(\frac{2\pi}{\sin \theta} \right)^{1/2} e^{\frac{i\pi}{4}} e^{-iz_1 \theta} \sum \nu_s^{1/2} \frac{e^{-i\tau_s z_1^{1/3} \theta}}{\beta_s}, \quad (12)$$

where β_* is written for the value of $\frac{\partial N}{\partial \nu}$ at $\nu=\nu_*$. Dividing this value of Π by the primary field (1) leads to the alternative formula

$$\frac{\Pi}{\Pi_{\text{prim}}} = \frac{2}{\bar{\alpha}} (\lambda D)^{1/3} e^{i \frac{3\pi}{4}} \sum_s \frac{1}{\beta_s} \exp \left[-i \tau_s \left(\frac{2\pi}{\lambda \bar{\alpha}^2} \right)^{1/3} D \right], \quad . \quad . \quad (13)$$

where the roots ν_s and the values β_s remain to be determined.

4. Comparison with Formulas for a Vertical Electric Dipole.

Before proceeding to these evaluations it may be of interest to compare the solution (13) with the corresponding solution for a vertical electric dipole. For the sake of brevity this will be called the electric case to distinguish it from the present, or magnetic case, and quantities in the electric case will be denoted by primed letters to distinguish them from the corresponding quantities in the magnetic case. Then the electric formula similar to (13), when expressed in the same units *, is

$$\frac{\Pi'}{\Pi'_{\text{prim}}} = \frac{2}{a} (\lambda D)^{1/2} e^{i \frac{3\pi}{4}} \sum_s \frac{1}{\beta_s} \exp \left[-i \tau_s \left(\frac{2\pi}{\lambda \bar{a}^2} \right)^{1/3} D \right], \quad (13')$$

where

$$\nu'_s = z_1 + \tau'_s z_1^{1/3}, \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (11')$$

ν'_s being the s^{th} root of the electric residue equation

$$N' \equiv \frac{\zeta'_{\nu-1/2}(z_1)}{\zeta_{\nu-1/2}(z_1)} - \frac{k_1 \psi'_{\nu-1/2}(z_2)}{k_2 \psi_{\nu-1/2}(z_2)} = 0, \quad . \quad . \quad . \quad (10')$$

and β'_s is the value of $\frac{\partial N'}{\partial \nu}$ at $\nu = \nu'_s$. The expression for N' differs from the corresponding value of N only in the inversion of the factor k_1/k_2 .

* Because of the changes in units and notation it would only lead to confusion to include cross-references to actual formulas in the papers of references (1) and (2). It should, however, be clearly understood that the electric formulas cited here are to be found in one or both of these papers.

5. Solutions of the Residue Equations.

In determining the values of the roots ν , either of the Debye expansions* for the Bessel functions ψ and ζ may be considered. These expansions are

$$\zeta_{\nu-1/2}(z_1) \cong \frac{2i}{z_1(\sin v_1)^{1/2}} \cos \left[\frac{\pi}{4} + \nu(\tan v_1 - v_1) \right], \quad . \quad . \quad . \quad (14)$$

or

$$\zeta_{\nu-1/2}(z_1) \cong \left(\frac{\pi}{6} \right)^{1/2} e^{-i\frac{\pi}{6}} \frac{\tan v_1}{z_1^{1/2}} e^{-i\nu\Phi} H_{1/3}^{(2)} \left(\frac{1}{3} \nu \tan^3 v_1 \right), \quad . \quad . \quad (15)$$

where

$$\cos v_1 = \frac{\nu}{z_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

and

$$\Phi = \tan v_1 - v_1 - \frac{1}{3} \tan^3 v_1.$$

Since z_2 is always larger than z_1 , only the first form of expansion need be considered for the function $\psi_{\nu-1/2}$, namely,

$$\psi_{\nu-1/2}(z_2) \cong \frac{1}{2z_2(\sin v_2)^{1/2}} \exp \left[-i\frac{\pi}{4} + i\nu(\tan v_2 - v_2) \right], \quad . \quad (17)$$

where

$$\cos v_2 = \frac{\nu}{z_2} = \frac{k_1}{k_2} \cos v_1. \quad . \quad . \quad . \quad . \quad . \quad (18)$$

On the assumption that the angle v_1 is small (verified *a posteriori*), the magnetic residue equation (10) reduces to one of the forms

$$\tan p \equiv \tan \left(\frac{\pi}{4} + \frac{1}{3} z_1 v_1^3 \right) = \frac{-ik_2}{k_1 v_1} \sin v_2, \quad . \quad . \quad . \quad (19)$$

or

$$\frac{e^{-i\frac{2\pi}{3}} H_{2/3}^{(2)} \left(\frac{1}{3} z_1 v_1^3 \right)}{H_{1/3}^{(2)} \left(\frac{1}{3} z_1 v_1^3 \right)} = \frac{ik_2}{k_1 v_1} \sin v_2, \quad . \quad . \quad . \quad (20)$$

depending on the use of expansions (14) and (15) respectively. The corresponding electric residue equation (10') becomes either

$$\tan p' \equiv \tan \left(\frac{\pi}{4} + \frac{1}{3} z_1 v_1'^3 \right) = -\frac{ik_1}{k_2 v_1'} \sin v_2', \quad . \quad . \quad . \quad (19')$$

or

$$\frac{e^{-i\frac{2\pi}{3}} H_{2/3}^{(2)} \left(\frac{1}{3} z_1 v_1'^3 \right)}{H_{1/3}^{(2)} \left(\frac{1}{3} z_1 v_1'^3 \right)} = \frac{ik_1}{k_2 v_1'} \sin v_2'. \quad . \quad . \quad . \quad (20')$$

* Cf. van der Pol and Bremmer, formulas (16 a) and (16 b).

In equations (20) and (20') the further substitutions

$$\frac{1}{3}z_1v_1^3 = \alpha e^{i\pi}, \quad \frac{1}{3}z_1v_1'^3 = \alpha' e^{i\pi}. \quad (21, 21')$$

lead finally to the forms

$$\frac{(3\alpha)^{1/3}[J_{2/3}(\alpha) - J_{-2/3}(\alpha)]}{J_{1/3}(\alpha) + J_{-1/3}(\alpha)} = \frac{e^{-i\frac{\pi}{3}}}{\delta}, \quad (22)$$

$$\frac{(3\alpha')^{1/3}[J_{2/3}(\alpha') - J_{-2/3}(\alpha')]}{J_{1/3}(\alpha') + J_{-1/3}(\alpha')} = \frac{e^{-i\frac{\pi}{3}}}{\delta'}, \quad (22')$$

where

$$\delta = \frac{-i}{z_1^{1/3}} \left(\frac{k_2^2}{k_1^2} - 1 \right)^{-1/2}, \quad (23)$$

$$\delta' = \frac{-i}{z_1^{1/3}} \frac{k_2^2}{k_1^2} \left(\frac{k_2^2}{k_1^2} - 1 \right)^{-1/2}. \quad (23')$$

Thus the difference in the solutions of the electric and magnetic cases depends primarily on the difference in the values of the earth-factors δ and δ' . But since k_2 is always greater than k_1 the presence of the factor $z_1^{-1/3}$ in the definition of δ shows that in the magnetic case $|\delta|$ is always small (at least for wave-lengths in the range from 1 to 10 metres in which we are interested), whereas $|\delta'|$ may vary from a small value (for a dielectric earth, that is, an earth of low conductivity) to an infinite value (for a perfectly conducting earth). Thus in equation (22) the right-hand side is always large and the solutions α will be very near the solutions of

$$J_{1/3}(\alpha) + J_{-1/3}(\alpha) = 0. \quad (24)$$

But the solutions of equation (22') will vary between values of α' near the solutions of equation (24) for a dielectric earth, and values of α' near the solutions of

$$J_{2/3}(\alpha') - J_{-2/3}(\alpha') = 0 \quad (24')$$

for an earth of high conductivity. It follows that the solutions in the magnetic case are always near those for a dielectric earth in the electric case, and further that the values of the earth constants are of much less importance in the magnetic case than in the electric case.

It may also be verified that for the magnetic case the solutions obtained from equation (19) and those from the more accurate equation (20) are practically indistinguishable *, so that only the simpler form (19) need be

* This has been pointed out by van der Pol and Bremmer in the recently published third part of their paper (reference⁽⁶⁾). They show that in the electric case the "tangent" form is sufficiently accurate for the case of dielectric earth, but that this is not true for the case of an earth of high conductivity.

considered. On introducing the parameters α and δ , therefore, the equation to be solved is

$$(3\alpha)^{1/3} \tan\left(\frac{\pi}{4} - \alpha\right) = -\frac{1}{\delta} e^{-i\frac{\pi}{3}}. \quad (25)$$

As the right-hand side of this equation is large the values of α_s must be near $(s + \frac{3}{4})\pi$. To a first approximation the actual solution may be written

$$\alpha_s = (s + \frac{3}{4})\pi - [3(s + \frac{3}{4})\pi]^{1/3} \delta e^{i\frac{\pi}{3}},$$

or, in terms of the original parameter τ_s ,

$$\tau_s = \bar{\tau}_s - \delta, \quad (26)$$

where

$$\bar{\tau}_s = \frac{1}{2}[3(s + \frac{3}{4})\pi]^{2/3} e^{-i\frac{\pi}{3}}. \quad (27)$$

This value $\bar{\tau}_s$ is identical with the approximate solution τ'_s given by van der Pol and Bremmer in the electric case for a dielectric (or strongly absorbing) earth. For an earth of high conductivity the corresponding solution in the electric case is more involved, but it has been fully treated by Wwedensky (reference⁽²⁾) and need not be considered here.

6. Determination of the Values of β_s .

The solution for the case of zero heights will be complete when the values of β_s have been determined. These are easily obtained by differentiating the appropriate expansion for N , namely,

$$N \cong -\sin v_1 \tan p - \frac{ik_2}{k_1} \sin v_2,$$

and remembering that at $v = v_s$ equation (19) is satisfied. Then the value of β_s is found to be

$$\beta_s \cong -\frac{k_2^2}{k_1^2} \sin^2 v_2 - \frac{ik_2}{k_1} \frac{\sin v_2}{z_1 v_1^2} + v_1^2, \quad v = v_s, \quad (28)$$

where the corresponding electric value is

$$\beta'_s \cong -\frac{k_1^2}{k_2^2} \sin^2 v'_2 - \frac{ik_1}{k_2} \frac{\sin v'_2}{z_1 v_1'^2} + v_1'^2, \quad v = v'_s. \quad (28')$$

Introducing the distance parameter η already used by van der Pol and Bremmer,

$$\eta = \frac{D}{a^{2/3} \lambda^{1/3}}, \quad (29)$$

the final formulas for the fields in the electric and magnetic cases are, respectively,

$$\frac{\Pi}{\Pi_{\text{prim}}} = 2 \left(\frac{\lambda}{a} \right)^{2/3} \eta^{1/2} e^{i \frac{3\pi}{4}} \sum_s \frac{1}{\beta_s} \exp[-i(2\pi)^{1/3} \tau_s \eta] \quad . \quad . \quad (30)$$

and

$$\frac{\Pi'}{\Pi'_{\text{prim}}} = 2 \left(\frac{\lambda}{a} \right)^{2/3} \eta^{1/2} e^{i \frac{3\pi}{4}} \sum_s \frac{1}{\beta'_s} \exp[-i(2\pi)^{1/3} \tau'_s \eta]. \quad . \quad . \quad (30')$$

If the distance at which the field is measured is large enough, one term of each of the series (30) and (30') will be sufficient, and if the primary fields in the two cases are adjusted to equal power output * the ratio of the Hertzian vectors in the magnetic and electric cases may be written in the very simple form

$$\frac{\Pi}{\Pi'} = \frac{1}{120\pi} \frac{\beta'_0}{\beta_0} \exp[-i(2\pi)^{1/3}(\tau_0 - \tau'_0)\eta]. \quad . \quad . \quad (31)$$

For a dielectric earth τ_0 and τ'_0 are so nearly equal that in this case the exponential factor could be omitted.

It should be noted that in the magnetic case, just as in the electric case, the ratio Π/Π_{prim} does not differ appreciably from the ratio of the electric fields, E/E_{prim} . This is easily verified by applying the field equation (3) for E_ϕ to the formulas (12) and (1) for Π and Π_{prim} . This gives, to a high degree of approximation in each case,

$$E_\phi = -i\omega\mu(-ik_1 a\Pi),$$

as long as the distance D at which the field is to be determined is large compared with the wave-length. Thus formulas in terms of the Hertzian vector ratio are directly applicable also to the field ratio, and it will be sufficient to write all formulas in terms of Π only.

* The power radiated by an electric current element in free space is

$$W' = \frac{1}{3} \pi Z \left(\frac{\Pi}{\lambda} \right)^2,$$

while that radiated by a magnetic current element is

$$W = \frac{1}{3} \pi Z \left(\frac{M}{\lambda} \right)^2.$$

Also the primary field in the electric case, corresponding to formula (1), is

$$\Pi'_{\text{prim}} = \frac{Z}{4\pi} \Pi \frac{e^{-ik_1 D}}{ik_1 D},$$

whence

$$\frac{\Pi_{\text{prim}}}{\Pi'_{\text{prim}}} = \frac{1}{Z^2} \frac{M}{\Pi} = \frac{1}{Z} \sqrt{\frac{W}{W'}} = \frac{1}{Z} \text{ when } W = W'.$$

7. Numerical Results for Zero Heights.

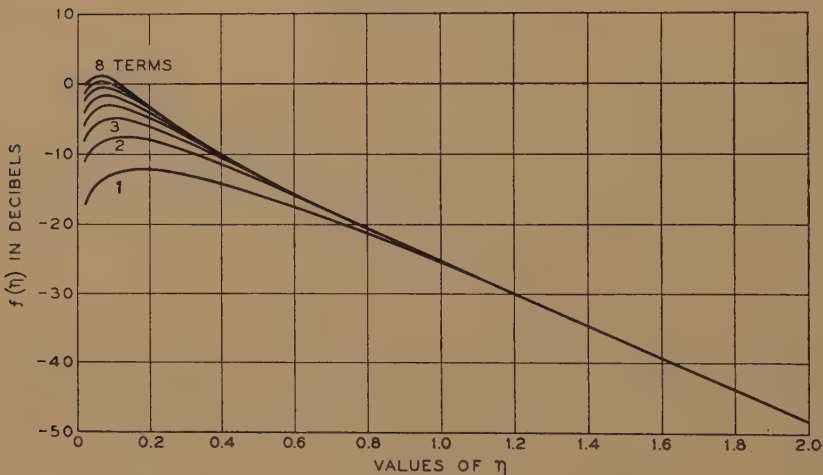
For purposes of numerical computation the formula (30) can be considerably simplified. In fact, in the expression (28) for β_s the first term

$$-\frac{k_2^2}{k_1^2} \sin^2 \nu_2 = 1 - \frac{k_2^2}{k_1^2}$$

is obviously much larger than either of the other terms, though the form (28) is interesting for purposes of comparison with the value (28') of β'_s . Retaining only this term in β_s makes its value independent of s , so that it can be taken outside the summation sign and the formula for the Hertzian vector ratio becomes

$$\frac{\Pi}{\Pi_{\text{prim}}} = 2(2\pi)^{2/3} e^{i\frac{3\pi}{4}} \delta^2 \eta^{1/2} e^{-i(2\pi)^{1/3} \delta \eta} \sum_s \exp[-i(2\pi)^{1/3} \bar{\tau}_s \eta]. \quad (32)$$

Fig. 2.



Modulus of the sum of the first eight terms of $f(\eta)$.

Or, finally, since the exponential term in δ does not differ appreciably from unity,

$$\frac{\Pi}{\Pi_{\text{prim}}} = 2(2\pi)^{2/3} e^{i\frac{3\pi}{4}} \delta^2 f(\eta), \quad (33)$$

where

$$f(\eta) = \eta^{1/2} \sum_s \exp[-i(2\pi)^{1/3} \bar{\tau}_s \eta]. \quad (34)$$

Thus the field on the surface at any distance η can be determined from the value of $f(\eta)$, which is completely independent of the earth constants, plus the value of the earth factor δ , which is easily determined when the appropriate values of the earth constants K and σ are known. In fig. 2

values of $|f(\eta)|$ have been plotted, in decibels, showing the effect of the first eight terms of the summation. For completeness, values of $|\delta|$ have been tabulated in Table I., for wave-lengths varying between 1 and 10 metres and for values of the effective radius between 1 and 3 times

TABLE I.
Values of the Earth-factor for Dielectric Earth
and for Sea-water.

$20 \log_{10} \delta .$					
Dielectric earth : $K=4, \sigma=10^{-3}$ mho/m.					
$\rho = \lambda$	1.	4/3.	3/2.	2.	3.
1....	-55.45	-56.29	-56.63	-57.46	-58.64
2....	-53.45	-54.28	-54.62	-55.46	-56.63
3....	-52.28	-53.11	-53.45	-54.29	-55.46
4....	-51.45	-52.29	-52.63	-53.46	-54.63
6....	-50.30	-51.13	-51.47	-52.30	-53.48
8....	-49.49	-50.32	-50.66	-51.49	-52.67
10....	-48.87	-49.70	-50.04	-50.85	-52.05
Sea-water : $K=80, \sigma=5$ mho/m.					
$\rho = \lambda$	1.	4/3.	3/2.	2.	3.
1....	-75.60	-76.43	-76.77	-77.61	-78.78
2....	-76.50	-77.33	-77.67	-78.50	-79.68
3....	-77.05	-77.89	-78.23	-79.06	-80.24
4....	-77.47	-78.30	-78.64	-79.48	-80.65
6....	-78.05	-78.88	-79.23	-80.06	-81.23
8....	-78.47	-79.30	-79.64	-80.47	-81.65
10....	-78.79	-79.62	-79.97	-80.79	-81.97

the actual radius, for the two cases of a dielectric earth ($K=4, \sigma=10^{-3}$ mho/m.) and of sea water ($K=80, \sigma=5$ mho/m.).

8. Effect of Raising Transmitter and Receiver above the Earth's Surface.

The effect of the height of the antennas on the field will be obtained by applying the method of transformation already used for the surface field, as given by the series (7), to the original series (6), which does not include

the assumption of zero heights. Carrying out the transformation it is found that, just as in the electric case, the height effect is obtained by multiplying each term of the transformed series (9) by two height factors, $g_s(h_1)$ and $g_s(h_2)$, of the form

$$g_s(h) = \left. \begin{aligned} \frac{\zeta_{v_s-1/2}(z_b)}{\zeta_{v_s-1/2}(z_a)}, \quad z_b = k_1 b = k_1(\bar{a} + h), \\ z_a \equiv z_1 = k_1 \bar{a}, \end{aligned} \right\} \dots \dots (34)$$

where h_1 and h_2 are the heights of transmitter and receiver respectively. Assuming that the heights are not so large that z_b is appreciably greater than z_a , the asymptotic expansions already used for the ζ -functions (equation (14)) will be applicable to each of the factors of $g_s(h)$, which therefore reduces immediately to

$$g_s(h) \cong \left(\frac{\sin v_a}{\sin v_b} \right)^{1/2} \frac{\cos p_b}{\cos p_a}, \quad \dots \dots (35)$$

where

$$\left. \begin{aligned} \cos v_a = \frac{v_s}{z_a} \equiv \cos v_1, \quad \cos v_b = \frac{v_s}{z_b} = \frac{z_a}{z_b} \cos v_a, \\ p_a = \frac{\pi}{4} + \frac{1}{3} z_a v_a^3, \quad p_b = \frac{\pi}{4} + \frac{1}{3} z_b v_b^3. \end{aligned} \right\} \dots \dots (36)$$

Equation (35) is the form of height factor required in the electric case, but in the magnetic case several simplifications are possible. Thus it has already been shown that $\tan p_a$ is always large, so that p_a has a value near $-(s + \frac{1}{2})\pi$, and in fact, to a high degree of approximation

$$p_a = -(s + \frac{1}{2})\pi + (-2\bar{\tau}_s)^{1/2} \delta, \quad \dots \dots (37)$$

where $\bar{\tau}_s$ is the parameter defined in equation (27). It follows that

$$\cos p_a = (-1)^s (-2\bar{\tau}_s)^{1/2} \delta. \quad \dots \dots (38)$$

Further the angle v_b is not large, so that a good approximation for v_b can be obtained from the definition (36), namely

$$v_b = v_a \Lambda_s^{1/2},$$

where

$$\Lambda_s = 1 - \frac{(2\pi)^{2/3}}{a^{1/3} \tau_s} \bar{h}, \quad \bar{h} = (\rho \lambda^2)^{-1/3} h, \quad \dots \dots (39)$$

and the height parameter \bar{h} is essentially equivalent to the height parameter $\lambda^{-2/3} h$ previously introduced by Eckersley⁽⁷⁾. It should also be noted that in the definition of Λ_s , $\bar{\tau}_s$ may be used for τ_s except at very small heights. Introducing Λ_s into the expression for p_b :

$$p_b = \frac{\pi}{4} + \frac{1}{3} (-2\tau_s)^{3/2} \Lambda_s^{3/2}; \quad \dots \dots (40)$$

the final form for $g_s(\bar{h})$ is

$$g_s(\bar{h}) = \frac{(-1)^s \cos \left[\frac{\pi}{4} + \frac{1}{3} (-2\tau_s)^{3/2} A_s^{3/2} \right]}{A_s^{1/4} (-2\bar{\tau}_s)^{1/2} \delta} \dots \dots \dots (41)$$

A further substitution,

$$g_s(\bar{h}) = \frac{1}{\delta} G_s(\bar{h}), \dots \dots \dots (42)$$

leads to a formula for the field with both antennas above the surface, corresponding to equation (30) for the field on the surface,

$$\frac{\Pi}{\Pi_{\text{prim}}} = 2(2\pi)^{2/3} e^{i\frac{3\pi}{4}} \eta^{1/2} \sum_s G_s(\bar{h}_1) G_s(\bar{h}_2) \exp [-i(2\pi)^{1/3} \bar{\tau}_s \eta], \dots (43)$$

where

$$G_s(\bar{h}) = \frac{(-1)^s \cos \left[\frac{\pi}{4} + \frac{1}{3} (-2\tau_s)^{3/2} A_s^{3/2} \right]}{A_s^{1/4} (-2\bar{\tau}_s)^{1/2}} \dots \dots \dots (44)$$

Equation (43) constitutes a complete solution of the problem of the magnetic dipole, and, except at very small heights, it is substantially independent of the values of the earth constants. If $\bar{\tau}_s$ is used for τ_s in the definition of $G_s(\bar{h})$ the solution for the case of a perfectly conducting earth is obtained. In this limiting case the field has a definite value as long as neither antenna is on the surface, but since

$$G_s(0) = \delta, \dots \dots \dots (45)$$

and δ is zero for a perfectly conducting earth, the field will vanish if either antenna is lowered to the surface.

9. Numerical Evaluation of the Effect of Height.

When the distances at which the field is to be measured are large, and the antennas are not too high above the surface, one term of the series (43) will be sufficient to determine the field. Thus the simplest formula for practical purposes is obtained by including only the term $s=0$:

$$\left| \frac{\Pi}{\Pi_{\text{prim}}} \right| = 2(2\pi)^{2/3} |f(\eta)| |G_0(\bar{h}_1) G_0(\bar{h}_2)| \dots \dots \dots (46)$$

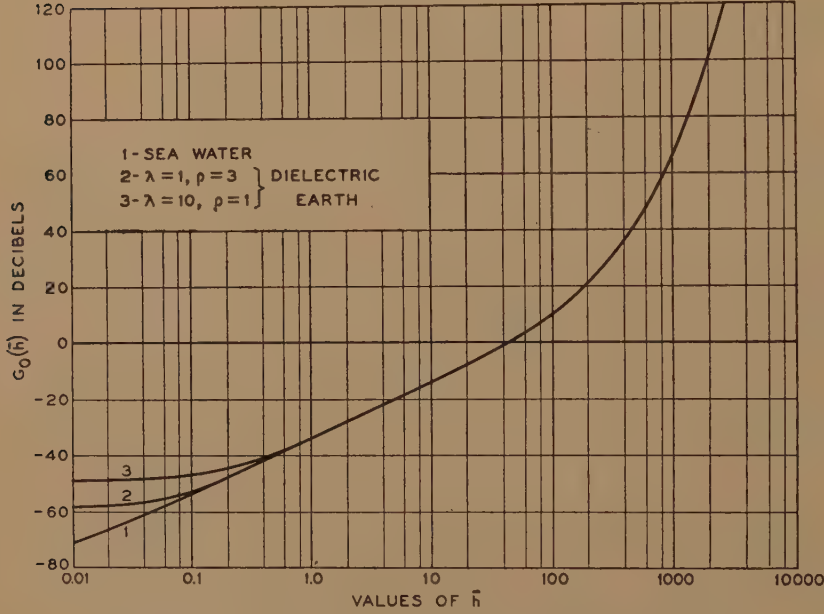
Also, since the height functions $G_0(\bar{h})$ are independent of the distance parameter η , the effects of height and distance can be evaluated separately. The values of $f(\eta)$ have already been shown in fig. 2. In fig. 3 the values of $G_0(\bar{h})$ are shown for heights ranging from about 1 cm. to 1000 m. From the two curves, values of $f(\eta)$, $G_0(\bar{h}_1)$ and $G_0(\bar{h}_2)$ can be read off and the

decibel values have simply to be added together to obtain the final field value.

Fig. 3 has been extended to very low heights to show how slight the effect of the earth constant is. The portion of the curve below $\bar{h}=1$ was drawn by approximating to the value of $\cos p_b$ in the formula for $G_0(\bar{h})$ in the same way as $\cos p_a$ was approximated in equation (38). Assuming that $\bar{h}/a^{1/3}$ is of the same order of magnitude as δ the approximation obtained is

$$\cos p_b = (-1)^s \frac{3(s + \frac{3}{4})\pi}{2\tau_s} \left[\delta + (2\pi)^{2/3} \frac{\bar{h}}{a^{1/3}} \right],$$

Fig. 3.



Height factor $G_0(\bar{h})$ for sea-water and dielectric earth.

from which it follows that

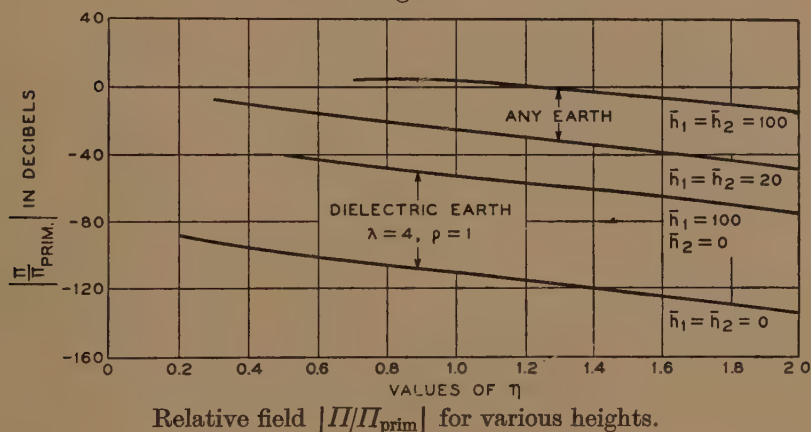
$$G_s(\bar{h}) = \delta + \frac{(2\pi)^{2/3} \bar{h}}{a^{1/3}}, \quad \bar{h} \leq 1. \quad \dots \quad (47)$$

Actually this approximation appears to be valid for values of \bar{h} up to about 10, but it is only for values below unity that the value of δ affects the result. The curve marked "sea water" in fig. 3 is substantially the same as the curve for a perfectly conducting earth, and is not appreciably affected by a change in wave-length within the range of 1 to 10 metres.

The curves for a dielectric earth show a slight divergence within this range at very low heights, so curves for two distinct values ($\lambda=1$, $\rho=3$, and $\lambda=10$, $\rho=1$) have been included.

Finally, in fig. 4, the complete formula (43) is shown graphically for various antenna heights. When both antennas are raised above the earth's surface the curves obtained are independent of the values of the earth constants and are thus directly applicable for transmission over any earth. When either or both antennas are on the earth, however, the field does depend on the earth constants, hence for definiteness the two lowest curves of fig. 4 have been drawn for a dielectric earth, using the value of δ corresponding to $\lambda=4$, $\rho=1$ in Table I. For any other value,

Fig. 4.



δ' , of the earth constant the field will be immediately determined by adding the algebraic difference ($\delta' - \delta$) to the decibel value of the field obtained from the curves.

In drawing the curves of fig. 4, the summation of the series (43) has been carried as far as $s=20$, and each curve has been started at the smallest value of η for which terms above $s=20$ are negligible. For distances beyond these minimum values, of course, fewer terms are required, but the curves as drawn give an accurate representation of the field of the magnetic dipole as expressed in equation (43).

References.

- (1) B. van der Pol and H. Bremmer, "The Diffraction of Electromagnetic Waves from an Electrical Point Source round a Finitely Conducting Sphere, with Applications to Radiotelegraphy and the Theory of the Rainbow," *Phil. Mag.* (7) xxiv. pp. 141-175 (July); and pp. 825-864 (Nov. Suppl. 1937).

- (2) B. Wwedensky, "The Diffractive Propagation of Radio Waves," *Tech. Phys. U.S.S.R.* ii. no. 6, pp. 624-639 (1935); iii. no. 11, pp. 915-925 (1936); and iv. no. 8, pp. 579-591 (1937).
- (3) G. N. Watson, "The Diffraction of Electric Waves by the Earth," *Roy. Soc. London, Proc. A*, xcv. pp. 83-99 (Oct. 1918); and pp. 546-563 (July 1919).
- (4) J. C. Schelleng, C. R. Burrows, and E. B. Ferrell, "Ultra-short-Wave Propagation," *Bell System Technical Journal*, xii. pp. 125-161 (Apr. 1933).
- (5) T. L. Eckersley and G. Millington, "Application of the Phase Integral Method to the Analysis of the Diffraction and Refraction of Wireless Waves round the Earth, *Roy. Soc. London, Phil. Trans.* cccxxvii. no. 778, pp. 273-309 (June 1938).
- (6) B. van der Pol and H. Bremmer, "The Propagation of Radio Waves over a Finitely Conducting Spherical Earth," *Phil. Mag.* (7) xxv. pp. 817-834 (June Suppl. 1938).
- (7) T. L. Eckersley, "Ultra-short-Wave Refraction and Diffraction," *I.E.E. Jl.* lxxx. pp. 286-304 (March 1937).

Bell Telephone Laboratories, Inc.
New York, N.Y.

XXXIX. *Direct Determination of Stresses from the Stress Equations in some Two-Dimensional Problems of Elasticity.*—Part II. *Thermal Stresses.*

By BIBHUTIBHUSAN SEN, Krishnagar College, Bengal *.

[Received May 3, 1938.]

IV. *Solution of Two-dimensional Thermo-elastic Equations.*

In continuation of the previous part †, we shall discuss in this paper a few problems of two-dimensional thermal stresses and show that these stresses can also be deduced directly from the stress equations and equations of compatibility. Thermal stresses are produced in a body by unequal distribution of temperature, which may be regarded as a specified function of coordinates and time. The distribution of temperature is determined by the laws of heat conduction, and there are two classes of problem according as the temperature is steady or variable.

If T denotes the temperature at the point (x, y) , and α the coefficient of linear thermal expansion (which is supposed to be constant), we have the following stress strain relations in plane problems of thermal stresses :

$$\left. \begin{aligned} e_{xx} &= \frac{1}{E} [(1+\sigma)\widehat{xx} - \sigma(\widehat{xx} + \widehat{yy} + \widehat{zz})] + \alpha T, \\ e_{yy} &= \frac{1}{E} [(1+\sigma)\widehat{yy} - \sigma(\widehat{xx} + \widehat{yy} + \widehat{zz})] + \alpha T, \\ e_{xy} &= \frac{2}{E} (1+\sigma)\widehat{xy}. \end{aligned} \right\} \dots \dots (4.1)$$

For plane strain problems we have also

$$e_{zz} = \frac{1}{E} [(1+\sigma)\widehat{zz} - \sigma(\widehat{xx} + \widehat{yy} + \widehat{zz})] + \alpha T = 0,$$

whence we get

$$\widehat{zz} = \sigma(\widehat{xx} + \widehat{yy}) - E\alpha T. \dots \dots (4.2)$$

* Communicated by the Author.

† “Direct Determination of Stresses from the Stress Equations in some Two-dimensional Problems of Elasticity.”—Part I. *Phil. Mag.* (7) xxvi. p. 98 (1938).

This result gives the modified values of strain components as follows :

$$\left. \begin{aligned} e_{xx} &= \frac{1+\sigma}{E} [\widehat{xx}(1-\sigma) - \sigma\widehat{yy} + E\alpha T], \\ e_{yy} &= \frac{1+\sigma}{E} [\widehat{yy}(1-\sigma) - \sigma\widehat{xx} + E\alpha T], \\ e_{xy} &= \frac{2(1+\sigma)}{E} \widehat{xy}. \end{aligned} \right\} \dots \dots (4.3)$$

Hence from the equations of equilibrium

$$\left. \begin{aligned} \frac{\partial \widehat{xx}}{\partial x} + \frac{\partial \widehat{xy}}{\partial y} &= 0, \\ \frac{\partial \widehat{xy}}{\partial x} + \frac{\partial \widehat{yy}}{\partial y} &= 0, \end{aligned} \right\} \dots \dots (4.4)$$

and the equation of compatibility

$$\frac{\partial^2 e_{xx}}{\partial y^2} + \frac{\partial^2 e_{yy}}{\partial x^2} = \frac{\partial^2 e_{xy}}{\partial x \partial y},$$

we get

$$\nabla_1^2 \left[\odot + \frac{E\alpha T}{1-\sigma} \right] = 0, \dots \dots (4.5)$$

where

$$\odot = \widehat{xx} + \widehat{yy}$$

and

$$\nabla_1^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \left\} \dots \dots (4.6)$$

As a solution of (4.5) we can put

$$\odot = \odot_T + \odot_0, \dots \dots (4.7)$$

such that

$$\odot_T = -\frac{E\alpha T}{1-\sigma}, \dots \dots (4.8)$$

and \odot_0 is a plane harmonic function satisfying the equation

$$\nabla_1^2 \odot_0 = 0. \dots \dots (4.9)$$

It is thus apparent that \odot_T is the contribution of thermal stresses to \odot , while \odot_0 , which is independent of the temperature, gives us stresses when there is no thermal expansion. In working out problems we shall first find out the stresses produced by thermal expansion. This expansion will no doubt give rise to certain stresses on the boundary of the solid. We can make the boundary free from stress by the addition of extra terms obtained on the hypothesis that there is no temperature distribution. We shall denote the stresses due to thermal expansion by the suffix T, while the suffix 0 will indicate the stresses deduced on the hypothesis that there is no thermal expansion.

The value of $\Sigma \odot_v$ can be found from (4.14) and (4.11) when T is given as a function of coordinates and time. Hence the values of $(xx)_T$, $(yy)_T$, and $(xy)_T$ can be written down at once from the above relations.

V. Solution suitable for a Semi-infinite Solid bounded by the Plane $x=0$.

When there is no thermal expansion the stresses $(\widehat{xx})_0$, $(\widehat{yy})_0$, and $(xy)_0$ are obtained as

$$\left. \begin{aligned} (\widehat{xx})_0 &= -\frac{1}{2}x \frac{\partial \odot_0}{\partial x} + \phi_1(x, y), \\ (\widehat{xy})_0 &= -\frac{1}{2}x \frac{\partial \odot_0}{\partial y} + \phi_2(x, y), \\ (\widehat{yy})_0 &= \odot_0 + \frac{1}{2}x \frac{\partial \odot_0}{\partial x} - \phi_1(x, y), \end{aligned} \right\} \dots \dots \dots (5.1)$$

where $\phi_1(x, y)$ and $\phi_2(x, y)$ are plane harmonic functions. If, in the case of non-steady distribution of temperature, the boundary $x=0$ be free from stress, we must have from (4.16) and (5.1)

$$\left. \begin{aligned} [\phi_1(x, y)]_{x=0} &= \left[\Sigma \frac{1}{\beta_v^2} \frac{\partial^2 \odot_v}{\partial y^2} \right]_{x=0} = p(y) \\ [\phi_2(x, y)]_{x=0} &= \left[-\Sigma \frac{1}{\beta_v^2} \frac{\partial^2 \odot_v}{\partial x \partial y} \right]_{x=0} = q(y) \end{aligned} \right\} \dots \dots \dots (5.2)$$

Determining $\phi_1(x, y)$ and $\phi_2(x, y)$ from the above relations and putting

$$\left. \begin{aligned} \int_{-\infty}^{\infty} p(\lambda) \log R \, d\lambda &= V_1, \\ \int_{-\infty}^{\infty} q(\lambda) \log R \, d\lambda &= V_2, \end{aligned} \right\} \dots \dots \dots (5.3)$$

where

$$R^2 = (\lambda - y)^2 + x^2,$$

we have from the equations of equilibrium (4.4)

$$\odot_0 = \frac{2}{\pi} \left[\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} \right]. \dots \dots \dots (5.4)$$

Using this value of \odot_0 and the values of $\phi_1(x, y)$ and $\phi_2(x, y)$ already obtained in (5.2), we obtain the stresses $(\widehat{xx})_0$, $(\widehat{xy})_0$, and $(\widehat{yy})_0$ from (5.1). Then the resultant stress components are given by

$$\left. \begin{aligned} \widehat{xx} &= (\widehat{xx})_T + (\widehat{xx})_0, \\ \widehat{xy} &= (\widehat{xy})_T + (\widehat{xy})_0, \\ \widehat{yy} &= (\widehat{yy})_T + (\widehat{yy})_0. \end{aligned} \right\} \dots \dots \dots (5.5)$$

VI. Thermal Stresses in a Circular Cylinder.

From the theory of heat conduction we know that when the temperature is not symmetrical about the axis we have

$$T = \sum_{v=1}^{\infty} \sum_{n=0}^{\infty} e^{-k\beta v^2 t} J_n(\beta_v r) [A_{n,v} \cos n\theta + B_{n,v} \sin n\theta], \quad (6.1)$$

in which $A_{n,v}$ and $B_{n,v}$ are constants depending on the initial values of the temperature.

Taking for example

$$T = \sum_{v=1}^{\infty} \sum_{n=0}^{\infty} A_{n,v} e^{-k\beta v^2 t} J_n(\beta_v r) \cos n\theta, \quad (6.2)$$

we have from (4.14)

$$\odot_v = -\frac{E\alpha}{1-\sigma} \sum_{n=0}^{\infty} A_{n,v} e^{-k\beta v^2 t} J_n(\beta_v r) \cos n\theta. \quad (6.3)$$

Since from (4.16) we have on transformation

$$\left. \begin{aligned} (\widehat{rr})_T &= -\sum_v \frac{1}{\beta_v^2} \left[\frac{1}{r^2} \frac{\partial^2 \odot_v}{\partial \theta^2} + \frac{1}{r} \frac{\partial \odot_v}{\partial r} \right], \\ (\widehat{\theta\theta})_T &= -\sum_v \frac{1}{\beta_v^2} \frac{\partial^2 \odot_v}{\partial r^2}, \\ (\widehat{r\theta})_T &= \sum_v \frac{1}{\beta_v^2} \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial \odot_v}{\partial \theta} \right), \end{aligned} \right\} \dots \dots \dots (6.4)$$

these stress components are given by

$$\begin{aligned} (\widehat{rr})_T &= -\frac{E\alpha}{1-\sigma} \sum_{v=1}^{\infty} \sum_{n=0}^{\infty} A_{n,v} \frac{e^{-k\beta v^2 t}}{\beta_v^2 r^2} [n(n-1)J_n(\beta_v r) + \beta_v r J_{n+1}(\beta_v r)] \cos n\theta, \\ (\widehat{r\theta})_T &= -\frac{E\alpha}{1-\sigma} \sum_{v=1}^{\infty} \sum_{n=0}^{\infty} n A_{n,v} \frac{e^{-k\beta v^2 t}}{\beta_v^2 r^2} [-(n-1)J_n(\beta_v r) + \beta_v r J_{n+1}(\beta_v r)] \sin n\theta, \\ (\widehat{\theta\theta})_T &= \frac{E\alpha}{1-\sigma} \sum_{v=1}^{\infty} \sum_{n=0}^{\infty} A_{n,v} J_n''(\beta_v r) \cos n\theta. \quad (6.5) \end{aligned}$$

On the boundary $r=a$, we have

$$\left. \begin{aligned} [(\widehat{rr})_T]_{r=a} &= -\sum_{v=1}^{\infty} \sum_{n=0}^{\infty} P_{n,v} \cos n\theta, \\ [(\widehat{r\theta})_T]_{r=a} &= -\sum_{v=1}^{\infty} \sum_{n=0}^{\infty} Q_{n,v} \sin n\theta, \end{aligned} \right\} \dots \dots \dots (6.6)$$

in which $P_{n,v}$ and $Q_{n,v}$ are known functions of a .

By applying the method given in 2 (a) of the previous part, or otherwise,

we get the stresses deduced from \odot_0 , which satisfies the equation $\nabla_1^2 \odot_0 = 0$. These are

$$\begin{aligned} (\widehat{rr})_0 &= \sum_{n=0}^{\infty} \left[-n(n-1)L_n \left(\frac{r}{a}\right)^{n-2} - (n+1)(n-2)M_n \left(\frac{r}{a}\right)^n \right] \cos n\theta, \\ (\widehat{r\theta})_0 &= \sum_{n=0}^{\infty} \left[n(n-1)L_n \left(\frac{r}{a}\right)^{n-2} + n(n+1)M_n \left(\frac{r}{a}\right)^n \right] \sin n\theta, \\ (\widehat{\theta\theta})_0 &= \sum_{n=0}^{\infty} \left[n(n-1)L_n \left(\frac{r}{a}\right)^{n-2} + (n+1)(n+2)M_n \left(\frac{r}{a}\right)^n \right] \cos n\theta. \end{aligned} \quad \dots \dots (6.7)$$

On the free boundary $r=a$, we must have

$$\left. \begin{aligned} (\widehat{rr})_0 + (\widehat{rr})_T &= 0, \\ (\widehat{r\theta})_0 + (\widehat{r\theta})_T &= 0. \end{aligned} \right\} \dots \dots \dots (6.8)$$

These results are satisfied if we take

$$\left. \begin{aligned} L_0 &= L_1 = 0, \\ M_0 &= \sum_{v=1}^{\infty} \frac{P_{0v}}{2}, \\ M_1 &= \sum_{v=1}^{\infty} \frac{Q_{1v}}{2} = \sum_{v=1}^{\infty} \frac{P_{1v}}{2}, \\ \text{and for } n \geq 2 \\ L_n &= - \sum_{v=1}^{\infty} \frac{nP_{nv} + (n-2)Q_{nv}}{2n(n-1)}, \\ M_n &= \sum_{v=1}^{\infty} \frac{P_{nv} + Q_{nv}}{2(n+1)}. \end{aligned} \right\} \dots \dots \dots (6.9)$$

VII. Solution for an Arbitrary Temperature.

When the temperature T has an arbitrary value $T(x, y)$ not governed by the law of heat conduction, the analysis given above will not be applicable. In this case on using the relation

$$\odot_T = - \frac{E\alpha}{1-\sigma} T,$$

in the equations of equilibrium (4.4) we get

$$\left. \begin{aligned} \nabla_1^2 (\widehat{xx})_T &= \frac{\partial^2 \odot_T}{\partial y^2} = - \frac{E\alpha}{1-\sigma} \frac{\partial^2 T}{\partial y^2} = U_1(x, y), \text{ (say)} \\ \nabla_1^2 (\widehat{xy})_T &= - \frac{\partial^2 \odot_T}{\partial x \partial y} = \frac{E\alpha}{1-\sigma} \frac{\partial^2 T}{\partial x \partial y} = U_2(x, y), \text{ (say)} \\ \nabla_1^2 (\widehat{yy})_T &= \frac{\partial^2 \odot_T}{\partial x^2} = - \frac{E\alpha}{1-\sigma} \frac{\partial^2 T}{\partial x^2} = U_3(x, y), \text{ (say)} \end{aligned} \right\} \dots \dots (7.1)$$

Then from the theory of potential we obtain

$$\left. \begin{aligned} (\widehat{xx})_T &= \frac{1}{2\pi} \iint U_1(\xi, \eta) \log r' d\xi d\eta, \\ (\widehat{xy})_T &= \frac{1}{2\pi} \iint U_2(\xi, \eta) \log r' d\xi d\eta, \\ (\widehat{yy})_T &= \frac{1}{2\pi} \iint U_3(\xi, \eta) \log r' d\xi d\eta, \end{aligned} \right\} \dots \dots \dots (7.2)$$

where $r' = [(x - \xi)^2 + (y - \eta)^2]^{\frac{1}{2}}$,
and the integration extends over the $\xi - \eta$ plane of the section.

VIII. Symmetrical Temperature Distribution.

Two-dimensional stresses, due to distribution of temperature which is symmetrical about an axis, can be easily deduced from the stress equations expressed in polar coordinates. Though this problem has been attacked in a different manner by previous investigators *, we shall briefly discuss it in this paper and deduce the expressions for stresses by a new method.

If we take the axis of z to be the axis of symmetry and the temperature to be a function of r only, we can assume that, due to thermal expansion, all stress components excepting $(\widehat{rr})_T$ and $(\widehat{\theta\theta})_T$ are zero. Since these stresses are also independent of θ , we have in polar coordinates the equation of equilibrium

$$\frac{\partial(\widehat{rr})_T}{\partial r} + \frac{(\widehat{rr})_T - (\widehat{\theta\theta})_T}{r} = 0. \dots \dots \dots (8.1)$$

We have also from (4.8)

$$\odot_T = (\widehat{rr})_T + (\widehat{\theta\theta})_T = -\frac{E\alpha T}{1-\sigma} \quad \text{or} \quad (\widehat{\theta\theta})_T = -\frac{E\alpha T}{1-\sigma} - (\widehat{rr})_T.$$

Substituting this value of $(\widehat{\theta\theta})_T$ in (8.1), we get

$$\frac{d}{dr} [r^2 (\widehat{rr})_T] = -\frac{E\alpha}{1-\sigma} Tr,$$

whence we obtain

$$\left. \begin{aligned} (\widehat{rr})_T &= -\frac{E\alpha}{1-\sigma} \frac{1}{r^2} \int Tr dr, \\ (\widehat{\theta\theta})_T &= -\frac{E\alpha}{1-\sigma} \left[T - \frac{1}{r^2} \int Tr dr \right] \end{aligned} \right\} \dots \dots \dots (8.2)$$

and

When there is no thermal expansion we have from (4.9)

$$\nabla_1^2 \odot = 0.$$

* *Vide* 'Theory of Elasticity,' by S. Timoshenko, art. 114.

For symmetrical distribution of stresses, the solution of this equation is

$$\odot_0 = C + D \log r, \quad \dots \dots \dots (8.3)$$

where C and D are constants.

A particular solution of the above, suitable for a solid cylinder, is

$$\odot_0 = C. \quad \dots \dots \dots (8.4)$$

Putting

$$(\widehat{\theta\theta})_0 = C - (\widehat{rr})_0$$

in the equation of equilibrium

$$\frac{\partial(\widehat{rr})_0}{\partial r} + \frac{(rr)_0 - (\widehat{\theta\theta})_0}{r} = 0,$$

we get

$$\frac{\partial}{\partial r} [r^2(\widehat{rr})_0] = Cr,$$

which gives

$$(\widehat{rr})_0 = \frac{C}{2} + \frac{E}{r^2},$$

E being a constant. For a complete cylinder E should be omitted, so that

$$\left. \begin{aligned} \widehat{rr} &= (\widehat{rr})_0 + (\widehat{rr})_T \\ &= \frac{C}{2} - \frac{E\alpha}{1-\sigma} \frac{1}{r^2} \int Tr \, dr, \text{ and} \\ \widehat{\theta\theta} &= (\widehat{\theta\theta})_0 + (\widehat{\theta\theta})_T \\ &= \frac{C}{2} - \frac{E\alpha}{1-\sigma} \left[T - \frac{1}{r^2} \int Tr \, dr \right] \end{aligned} \right\} \dots \dots \dots (8.5)$$

The value of the constant C can be easily obtained from the boundary condition. Proceeding in the same way we can find the thermal stresses in a solid bounded by concentric cylindrical surfaces due to symmetrical distribution of temperature.

Conclusion.

In this paper two-dimensional thermal stresses are obtained directly from the stress equations of equilibrium and compatibility in certain problems of plane deformation. We find that $\odot (= \widehat{xx} + \widehat{yy})$ can be split up into two parts, one \odot_T depending on temperature and the other \odot_0 independent of it. Stresses are deduced from these two parts of \odot , and the constants involved are so adjusted as to satisfy the required boundary conditions. In many problems stresses contributed by \odot_0 are already well known. Hence in these cases only the stresses produced by \odot_T are to be found. When the temperature is variable \odot_T is at once given by (4.8) and the values of the corresponding stress components can be easily obtained as in (4.16) in terms of the above.

Krishnagar College,
Bengal.

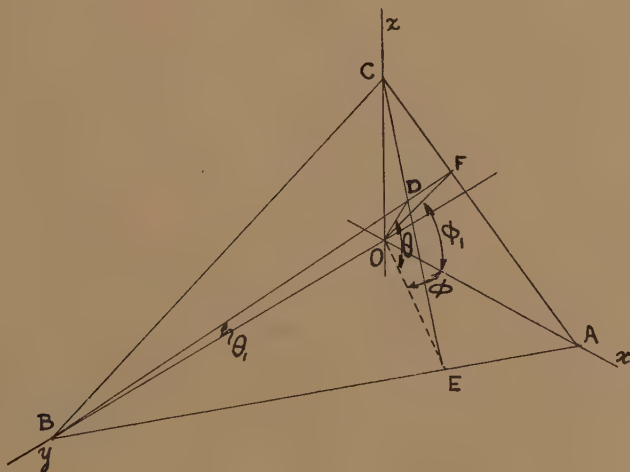
XL. *A Graphical Construction for Stress.*

By JAMES J. GUEST *.

[Received October 24, 1938.]

As increasing attention is being paid to complex cases of stress, to meet the preference which many have for graphical methods, I propose the following as a simple method of determining the stress on any plane when the principal stresses are known.

Fig. 1.



If the principal stresses in the directions of axes Ox_1 , Oy_1 , Oz_1 be ϖ_1 , ϖ_2 and ϖ_3 , and it is required to ascertain the stress on a plane ABC (fig. 1) given by the azimuth $AOE = \phi$ and altitude $EOD = \theta$ of its normal OD, where E is the point in which CDE meets AB, the determination will be complete when the normal component f_n and the shear components s_1 and s_2 , along and perpendicular to CDE, have been found.

The stress system is regarded as composed of a volumetric stress equal to one of the principal stresses, say ϖ_3 , superposed upon a stress system $p_1 = \varpi_1 - \varpi_3$, $p_2 = \varpi_2 - \varpi_3$, $p_3 = 0$. The stress due to the latter system is first considered.

* Communicated by the Author.

Along a line OGK (fig. 2) set off $OG = \frac{1}{2}(p_1 + p_2)$ and then $GH = \frac{1}{2}(p_1 - p_2)$ at an angle $HGK = 2\phi$. Draw OL at an angle $GOL = \theta$ and then HK perpendicular to OGK , KL to OL , LM to OG , and HV to KL . Finally mark off $MU = \varpi_3$.

Then $OU = f_n$, $LM = s_1$ and $KV = s_2$.

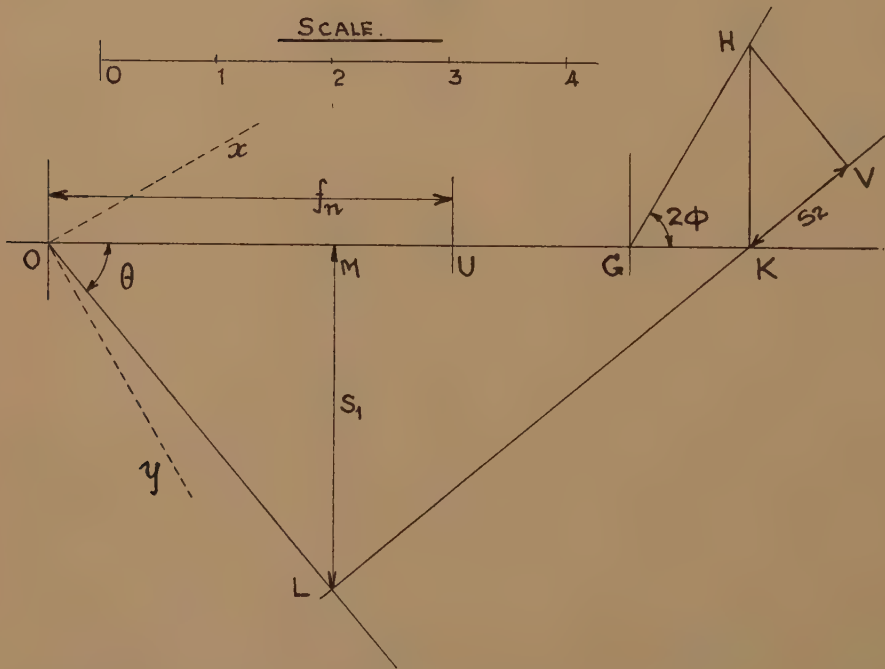
For the normal stress due to p_1 and p_2 is

$$p_1 \cos^2 \phi \cos^2 \theta + p_2 \sin^2 \phi \cos^2 \theta = \frac{1}{2}(p_1 + p_2 + p_1 - p_2 \cos 2\phi) \cos^2 \theta \\ = (OG + GK) \cos^2 \theta = OK \cos^2 \theta = OM,$$

so that

$$f_n = OM + \varpi_3 = OU.$$

Fig. 2.



The shear stress along CDE , s_1 , is similarly

$$(p_1 \cos^2 \phi + p_2 \sin^2 \phi) \cos \theta \cos \phi = OK \sin \theta \cos \theta = LM,$$

and that perpendicular to CDE , s_2 , is

$$p_1 \cos \theta \cos \phi \sin \phi - p_2 \cos \theta \sin \phi \cos \phi = (p_1 - p_2) \sin \phi \cos \phi \cos \theta \\ = GH \sin 2\phi \cos \theta = HK \cos \theta = KV.$$

As illustration suppose that, the principal stresses being $\varpi_1 = 8$, $\varpi_2 = 4$, and $\varpi_3 = 1$, the stress on a plane for which $\phi = 30^\circ$ and $\theta = 50^\circ$ be required.

If ϖ_3 be selected as the intermediate stress $p_1=7$ and $p_2=3$, so that $OG=\frac{1}{2}(p_1+p_2)=5$ and $GH=\frac{1}{2}(p_1-p_2)=2$ and the figure is easily constructed (fig. 2) to scale.

If the orientation of the plane ABC upon which the stress is required is given in another manner, the necessary angles are to be found from the geometry of the figure OABC. Thus, suppose that $\varpi_2=4$ be selected as the volumetric stress, so that it is necessary to find the azimuth and altitude of OD when OB is the pole. The construction would be as in

Fig. 3.

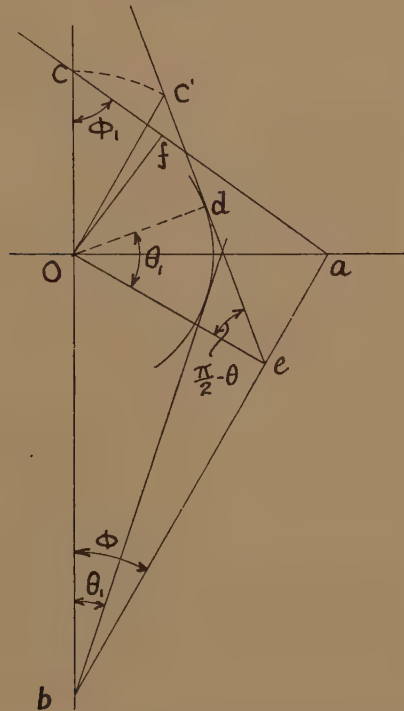


fig. 3, where Oab is a plane, ab being drawn at an angle $Oba=\phi$, Oe is drawn perpendicular to ab and then edc at an angle $\frac{\pi}{2}-\theta$ to Oe . Drawing Oc' perpendicular to Oe and Od perpendicular to ec' gives the rabated triangle $Oedc'$, so that, making $OC=Oc'$ the shape OAC is obtained, giving the new azimuth ϕ_1 . Drawing a circle, centre O and radius Od , and a tangent to it bd' , gives the new altitude Obd' or θ_1 .

Using ϕ_1 and θ_1 the diagram, fig. 4, can be drawn. The normal stress will be the same as that obtained in fig. 2, as will the resulting shear stress in the two cases.

XLI. *Some Problems of Finite Strain.*—II.

By B. R. SETH, M.A., D.Sc.*

[Received July 26, 1938.]

IN this part we shall deal with the case of a spherical shell turned inside out—a problem which cannot be satisfactorily solved with the help of the ordinary theory. The case of a cylindrical tube turned inside out has been dealt with elsewhere †. Except near the edges of the shell the displacement may be taken as symmetrical in all directions. Hence we can put

$$u = (1 - \beta)x, \quad v = (1 - \beta)y, \quad w = (1 - \beta)z, \quad \dots \quad (1)$$

where β is a function of $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$ only.

The stress components are given by

$$\widehat{xx} = \frac{1}{2}\lambda[3(1 - \beta^2) - (r^2\beta'^2 + 2r\beta\beta')] + \mu \left[(1 - \beta^2) - x^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right) \right], \quad (2.1)$$

$$\widehat{yy} = \frac{1}{2}\lambda[3(1 - \beta^2) - (r^2\beta'^2 + 2r\beta\beta')] + \mu \left[(1 - \beta^2) - y^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right) \right], \quad (2.2)$$

$$\widehat{zz} = \frac{1}{2}\lambda[3(1 - \beta^2) - (r^2\beta'^2 + 2r\beta\beta')] + \mu \left[(1 - \beta^2) - z^2 \left(\beta'^2 + \frac{2\beta\beta'}{r} \right) \right], \quad (2.3)$$

$$\widehat{yz} = -\mu yz \left(\beta'^2 + \frac{2\beta\beta'}{r} \right), \quad \dots \quad (2.4)$$

$$\widehat{zx} = -\mu zx \left(\beta'^2 + \frac{2\beta\beta'}{r} \right), \quad \dots \quad (2.5)$$

$$\widehat{xy} = -\mu xy \left(\beta'^2 + \frac{2\beta\beta'}{r} \right), \quad \dots \quad (2.6)$$

where $\beta' = d\beta/dr$.

Each body stress equation for no body force leads to the same differential equation for β , viz.,

$$4(\lambda + 2\mu)\beta\beta' + 2r(\lambda + 3\mu)\beta'^2 + (\lambda + 2\mu)(r^2\beta' + r\beta)\beta'' = 0. \quad \dots \quad (3)$$

Putting

$$r\beta' = \beta V, \quad \frac{\lambda + 3\mu}{\lambda + 2\mu} = f,$$

we get

$$(1 + V)\beta \frac{dV}{d\beta} + V^2 + 2fV + 3 = 0, \quad \dots \quad (3.1)$$

* Communicated by the Author.

† Shepherd and Seth, Proc. Roy. Soc. A, clvi. pp. 186–190 (1936).

which gives

$$\log \beta = \frac{f-1}{\sqrt{3-f^2}} \tan^{-1} \left\{ \frac{V+f}{\sqrt{(3-f^2)}} \right\} - \frac{1}{2} \log (V^2 + 2fV + 3) + C, \quad (3.2)$$

where C is a constant of integration.

Since

$$f = \frac{\lambda + 3\mu}{\lambda + 2\mu} = \frac{3-4\eta}{2(1-\eta)},$$

where η is Poisson's ratio, and since $0 < \eta < \frac{1}{2}$, we see that

$$1 < f < \frac{3}{2}.$$

Again we have

$$r \frac{dV}{dr} = \beta V \frac{dV}{d\beta}, \quad \dots \dots \dots (4)$$

from which, using (3.1), we get

$$\log r = \frac{1}{6} \log (V^2 + 2fV + 3) - \frac{1}{6} \log V^2 - \frac{3-f}{3\sqrt{(3-f^2)}} \tan^{-1} \left\{ \frac{V+f}{\sqrt{(3-f^2)}} \right\} + D, \quad (5)$$

where D is a constant.

Equations (3.2) and (5) give the relation between β and r in terms of the parameter V. C and D are to be determined from the boundary conditions.

If r' is the distance of a point of the shell from the centre in the unstrained state, which is at a distance r in the strained state, then

$$r' = r - r(1 - \beta) = r\beta,$$

so that

$$\frac{dr'}{dr} = r\beta' + \beta = \beta(1 + V). \quad \dots \dots \dots (6)$$

In the ordinary case of a shell subjected to normal tractions dr'/dr is positive throughout, and hence $V > -1$. But in the case of the shell turned inside out dr'/dr is negative throughout, and hence $V < -1$.

Moreover, as an india-rubber shell can be easily turned inside out, we should take

$$\eta = 0.49, \quad (\lambda = 49\mu).$$

In spherical polar coordinates the non-vanishing stresses can be written as

$$\begin{aligned} \bar{r}\bar{r} &= \frac{1}{2}(3\lambda + 2\mu)(1 - \beta^2) - \frac{1}{2}(\lambda + 2\mu)(r^2\beta'^2 + 2r\beta\beta') \\ &= \frac{1}{2}(3\lambda + 2\mu)(1 - \beta^2) - \frac{1}{2}(\lambda + 2\mu)\beta^2(2V + V^2), \quad \dots \dots (7.1) \end{aligned}$$

$$\bar{\theta}\bar{\theta} = \bar{\phi}\bar{\phi} = \frac{1}{2}(3\lambda + 2\mu)(1 - \beta^2) - \frac{1}{2}\lambda\beta^2(2V + V^2). \quad \dots \dots (7.2)$$

Introducing a new symbol, R, we can write

$$R \equiv 1 - \frac{2\bar{r}\bar{r}}{3\lambda + 2\mu} = \frac{\beta^2}{1 + \eta} [2\eta + (1 - \eta)(1 + V)^2]. \quad \dots \dots (7.3)$$

Putting $\eta=0.49$, we can rewrite (3.2), (5), and (7.3) as

$$\log (A\beta)=(0.006081) \tan^{-1}(0.714214V+0.728218) - \frac{1}{2} \log (V^2+2.039216V+3), \quad . \quad . \quad . \quad (8.1)$$

$$\log (Br)=- (0.204759) \tan^{-1}(0.714214 V+0.728218) + \frac{1}{6} \log (V^2+2.039216 V+3)-\frac{1}{6} \log V^2, \quad . \quad . \quad . \quad (8.2)$$

$$\log (A^2R)=2 \log (A\beta)+\log [(1+V)^2+1.92157]-0.4656161. \quad . \quad . \quad (8.3)$$

In (8) all logarithms are to the base 10.

Table I. gives the corresponding values of V , $\log (A\beta)$, $\log (Br)$, and $\log (A^2R)$.

TABLE I.

$-V.$	$-\log (A\beta).$	$\log (Br).$	$-\log (A^2R).$
1	0.14613	0.04587	0.47422
1.5	0.17235	0.06576	0.47354
1.545	0.17696	0.06879	0.47346
1.8	0.20801	0.08733	0.47313
2.0	0.23652	0.10233	0.47305
2.2	0.26702	0.11687	0.47312
2.5	0.31407	0.13692	0.47346
3.0	0.39059	0.16484	0.47435
4.0	0.52446	0.20356	0.47624
5.0	0.63275	0.22781	0.47775
10.0	0.96712	0.27613	0.48119
∞	∞	0.32163	0.48472

Assuming that suitable forces have been applied to the edge to keep it spherical we see that the boundary conditions are that \widehat{rr} must vanish both on the inner and outer surface of the shell. If a, b be its inner and outer radii when it has been turned inside out,

$$\widehat{rr}=0 \text{ over } r=a, r=b(b>a).$$

This condition implies that an increment of $\log (b/a)$ in $\log (Br)$ corresponds to a zero increment in $\log (A^2R)$. Since the variation in the values of $-\log (A^2R)$ with V is small, we should not use the graphical method of plotting $-\log (A^2R)$ against $\log (Br)$. By the method of trial and error we find that $\log (A^2R)$ has the same value for $V=-1.545$ and $V=2.5$. We have taken only a particular case. Any number of cases can be built up with the help of (8).

If we take $a=1$, Table I. gives $b=1.1698$. The corresponding values of β are found to be

$$\beta_a=1.1476, \quad \beta_b=0.83687,$$

from which we obtain Table II.

The percentage increase in the external radius is about 2 per cent., and in the thickness it is about 0.07 per cent. These can be approximately verified by measurement, which, on account of the edge-effect, should be made at some distance from the edge.

The constants A and B are given by

$$\begin{aligned} 2 \log A &= -0.47346, \\ \log B &= 0.06879. \end{aligned}$$

TABLE II.

	Internal radius.	External radius.	Thickness.
Right side out	1	1.1698	0.1698
Inside out	0.9790	1.1476	0.1686

Table I. can be used to determine the value of R corresponding to any value of V. Having known R (7.3) gives \widehat{rr} . The relation between \widehat{rr} and $\widehat{\theta\theta}$, which is independent of V, is easily seen to be

$$(\lambda + 2\mu)\widehat{\theta\theta} = \mu(3\lambda + 2\mu)(1 - \beta^2) + \lambda\widehat{rr}. \quad (9)$$

In the present case we find

$$\frac{\widehat{rr}}{\mu} = \frac{149}{2}(1 - R), \quad (10.1)$$

$$\frac{\widehat{\theta\theta}}{\mu} = \frac{149}{51}(1 - \beta^2) + \frac{49}{51} \frac{\widehat{rr}}{\mu}. \quad (10.2)$$

The corresponding values of V, r , \widehat{rr} , and $\widehat{\theta\theta}$ are given in Table III.

TABLE III.

-V.	r .	$-\widehat{rr}/\mu$.	$-\widehat{\theta\theta}/\mu$.
1.545	1	0	0.9257
1.8	1.0436	0.0551	0.4661
2.0	1.0803	0.0708	0.0708
2.2	1.1171	0.0581	-0.3246
2.5	1.1698	0	-0.8748

Tables II. and III. give results one expects to get when an india-rubber shell is turned inside out. The thickness of the shell decreases, (\widehat{rr}/μ) does not vary very much with r , and the absolute value of $\widehat{\theta\theta}$ is greatest at the inner boundary.

Hindu College,
Delhi, India.

XLII. *Quantum Theory of Gravitation.*

By A. KALIVARIIS, D.Sc., O. A. Stathatos Research Fellow †.

[Received October 7, 1938.]

ABSTRACT.

The connexion of gravitational theory with quantum mechanics is here illustrated by applying the principle of the existence of a gravity current and charge.

The linear equations of the gravitational field are constructed relativistically throughout.

An expression for the gravitational constant γ is given by means of atomic constants as \hbar , c , N , and m_0 .

The interaction of both gravitational and electromagnetic fields is expounded in the frame of quantum mechanics, whereas the cosmological case is taken into account in analogy with the theory of atomic nuclei.

1.

PROF. G. H. LIVENS, Dr. I. Viney ⁽¹⁾, and collaborators introduced as a basic concept of their unitary theory of gravitation and electricity the existence of a gravity charge, associated with every particle of matter and possessing only one sign, either positive or negative. Such a gravitational charge on any elementary particle is, in fact, very small in comparison with the electric charge of that particle of matter; hence it has been pointed out by Prof. Livens that "a charge, however small, may still have a preponderating inertia if sufficiently condensed."

From another standpoint Bohr and Gamow ⁽²⁾ have recently referred the gravitational radiation to the small probability of emission of neutrinos.

We shall here approach the question of the nature of the gravitational force in agreement with Livens concepts, namely, by adopting the principle of the existence of a gravity charge and taking into account the mutual force of a system of electrons instead of two elementary particles.

The idea of the usefulness of an electronic assembly for explaining gravitation was emphasized by Prof. O. W. Richardson as early as 1916 in his book 'The Electron Theory of Matter' (Cambridge, pp. 607-624).

One might, roughly speaking, compare the classical theory of gravitation, which is characterized by the Newton's constant γ , of which we are aware,

† Communicated by the Author.

to the statistical theory with its Boltzmann's constant K . It thus may be said to discover a law for the gravitational field in the case of an assembly of particles recognized by a new constant either $\gamma \cdot N = Z$ in the potential equation or γN^2 in the law of attraction which will correspond to the gas constant $R = K N$; N in all cases stands for Avogadro's universal number ^(1, III).

Such a law of gravitation alone can be compatible with the quantum theory, at least if we wish at the outset to avoid using cosmological terms in our equations.

2. The Law of Attraction.

Let us regard the motion of one body with a mass m in the central field of force of a second fixed mass m . In the quantum theory the angular momentum $r \times p = n\hbar$, a constant of the motion, possesses the components

$$m_x = yp_z - zp_y, \quad m_y = zp_x - xp_z, \quad m_z = xp_y - yp_x,$$

and it commutes with the Hamiltonian of the system in the non-relativistic case, namely, with the function

$$H = 1/2m(p_x^2 + p_y^2 + p_z^2) + E_{\text{pot}};$$

E_{pot} is the potential energy, a function only of $r^2 = (x^2 + y^2 + z^2)$.

The attraction-energy of the gravitating bodies, each with a mass m , will then be

$$E_{\text{pot}} = -\gamma \cdot \frac{m^2}{r}; \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

m denotes the total mass of N electrons which should be contained in a gram atom :

$$m = Nm_0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

We restrict ourselves to masses of that value (2), and we regard this number of N electrons as being distributed inside a whole sphere of density $\mu = m/\Omega$; Ω is any volume ξ^3 and ξ the radius of the sphere. Let the second body of a mass (2) thus move in the neighbourhood of the sphere as a fixed centre over a distance r .

Their gravitational energy is therefore equal to

$$E_{\text{pot}} = -\frac{4\pi}{3} \gamma \cdot \frac{m\mu}{\Omega} \cdot r^2;$$

the attraction energy reaches its maximum value on the surface of the sphere, in which $r = \xi$, and the previous equation reads

$$E_{\text{pot}} = -\frac{4\pi}{3} \gamma \cdot \frac{m^2}{r}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Such a mechanical system must fulfil the conservation law, namely, the total energy of the system should be reckoned as

$$-W = E_{\text{pot}} + E_{\text{cin}} = -\frac{2\pi}{3} \gamma \cdot \frac{m^2}{r} = \text{const.} \quad (4)$$

3. The Emission of Gravitational Radiation.

Let us now see how to fit it in with the radiation theory. The rate of emission of radiation of frequency $\nu = (H - H')/h$, owing to a transition of the system in the state a from an energy-level H to another H' , less than H , of some state a' , will be proportional to the square of the modulus of the matrix element, associated with this transition :

$$l = \frac{4}{3} \cdot \frac{(2\pi\nu)^4}{c^3} |a| D |a'|^2 \text{ sec}^{-1}.$$

D is the electric moment of the system ; the above matrix element has to replace the Fourier amplitude of D .

One gets the transition probability per unit of time by dividing l by the energy $h\nu$ of that quantum of frequency ν_{ik}

$$l = \frac{64\pi^4}{3} \cdot \frac{e^2}{hc^3} \cdot \nu_{ik}^3 \cdot r_{ik}^2 \quad (5)$$

The reciprocal value $1/l$ of this probability gives us the life time t of the quantum. We may simplify the formula (5) by virtue of the quantum condition :

$$n\hbar = m_0 \cdot r_{ik}^2 \cdot \nu_{ik}, \quad (6)$$

and then we have

$$l = \frac{64\pi^4}{3} \cdot \frac{e^2}{m_0 c^3} n \nu_{ik}^2 = \frac{64\pi^4}{3} \cdot n t_e \cdot \nu_{ik}^2 \quad (7)$$

The law of change of the transition probability with time

$$i\hbar \frac{d\psi}{dt} = H\psi$$

leads to a constant energy, the total Hamiltonian H , itself being put in diagonal form of its matrix elements ; the eigenstates of the latter represent the stationary states of the system.

From this point onward we adopt the usual formalism of the quantum theory of electromagnetic radiation in order to apply it to the emission of gravitational radiation. Correspondingly, the probability of emission of a gravitational quantum by an electronic system such as that of equation (2) can be attained by combining the probability (7) for emission of light quanta with the total energy (4) of the gravitating masses :

$$W = -\frac{2\pi}{3} \gamma \cdot \frac{N^2 m_0^2}{r} = \hbar \cdot l = \frac{64\pi^4}{3} \cdot \frac{\hbar c}{\lambda_{ik}^2} \cdot n r_0 \quad (8)$$

That is to say, the probability of emission of gravitational quanta by our mechanical system is proportional to the quantity $\sqrt{\gamma} \cdot Nm_0$, analogous to the electric charge which produces light quanta.

In the right-hand side of equation (8) we find a whole multiple of the classical radius r_0 , which is known from scattering experiments in nuclear physics as the minimum distance ; we are therefore led to replace r in the left-hand side of (8) by nr_0 , and again, by using the wave mechanical relation $p = h/\lambda_{ik}$, to put down as resonance condition $\lambda_{ik} = \pi r = n\pi r_0$. Thus $\lambda_{ik} = n\lambda_0$, where λ_0 should be a minimum wave-length πr_0 .

The resulting simplifications will reduce (8) to

$$W = -\gamma \cdot \frac{N^2 m_0^2}{nr_0} = \frac{32\pi\hbar c}{nr_0}, \quad . \quad . \quad . \quad . \quad . \quad (9)$$

whence we deduce the gravitational constant

$$\gamma = \frac{32\pi\hbar c}{N^2 m_0^2} \text{ gm.}^{-1} \text{ cm.}^3 \text{ sec.}^{-2} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Inserting in (10) the most probable values of $\hbar = 6.547 \times 10^{-27}$, $c = 2.998 \times 10^{10}$, $m = 9.02 \times 10^{-28}$, and $N = 6.06 \times 10^{+23}$, we obtain the value

$$\gamma = 6.6 \times 10^{-8} \text{ C.G.S.,}$$

in perfect agreement with the observational value.

If we wish to accept the data given recently by Millikan ⁽⁴⁾, *i. e.*, $\hbar = 6.61 \times 10^{-27}$ and $N = 6.03 \times 10^{+23}$, we then get

$$\gamma = 6.7 \times 10^{-8} \text{ C.G.S.}$$

4 Proof of the Constant N ; the Cosmological Law.

We must here enquire why nature favours a gram-molecule of electrons, and not, for instance, a kilogram-molecule ; although such a foundation of the universal constant N can no longer result from applying notions of Newton's law of attraction, nevertheless this proof is quite possible if we require the use of the cosmological form of the gravitational force.

The said cosmological law tends to avoid infinities of the G-potential in cosmical regions by assuming the total mass of the universe to be uniformly distributed on the surface of a sphere with a definite world radius R ; this is taken care of by the exponential function

$$E = \zeta^2 \cdot \frac{e^{-kr}}{r}, \quad . \quad . \quad . \quad . \quad . \quad (11)$$

in which ζ means a constant with the dimension of a charge and k a very small reciprocal length $\sim 1/R_0$.

The large length R_0 will be disclosed by the following equality :

$$W = \frac{1}{2} \gamma \cdot \frac{m_0 m_p}{r_0} = \frac{e^2}{R_0}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

which implies, according to (10),

$$R_0 = \frac{N^2 \alpha}{32 \pi^2} \cdot \frac{m_0}{m_p} \cdot r_0 \sim 1.2 \times 10^{27} \text{ cm.} \quad . \quad . \quad . \quad . \quad (13)$$

Because $R_0 = ct_0$, we conclude about them that :

“ An assembly of N pairs of electrons has for its proper time the Hubble's origin of time $t_0 = 4 \times 10^{16}$ sec. expressed in the atomic unit of time $e^2/m_0 c^3$ apart from a constant factor $\alpha m_0 / 32 \pi^2 m_p$.”

There α means the fine structure constant and m_p the proton mass.

We must look upon the fact embodied in equations (11), (12), and (13) as showing why nature prefers a number of N pairs of electrons in order to give rise to a gravitation law compatible with the quantum theory.

Because the true laws of nature are those of quantum mechanics therefore a reconciliation of Newton's gravitation and Hubble's recession may be illustrated in a complete quantum theory of gravitation.

Without dealing with the question at all deeply we may remark that some arguments may be brought forward in support of this idea.

We thus add that the total mass of a fictitious universe with a radius R_0 can similarly be expressed in terms of the atomic constants α , N , m_0 , and m_p , in keeping with Mach's principle and with the help of a Milne's relation

$$M_0 = \frac{c^3 t_0}{\gamma}.$$

Another feature of the connexion between atomic theory and cosmology may be found if we ask for the mass density of the electronic assembly (2), which is

$$\mu = \frac{N m_0}{(N \alpha r_0)^3} \sim 10^{-31} \text{ gm./cm.}^3;$$

that is to say, μ has about the same order of magnitude as the average density of nebular matter as calculated by Shapley.

5. Linear Equations for the Gravitational Field.

We now proceed to the building up of linear equations for the gravitational field. We shall here consider not the G-field of vacuum which may correspond to the Maxwell field without electric charge, namely, for $\text{div } \mathbf{E} = 0$, but the general case of the gravitational field owing to matter.

As our starting-point for the G-field equations we take the principle conceived by Livens and Viney of the occurrence of a gravity charge

associated with every particle of matter. In agreement with their investigations, we shall state once more that in a quantum theory of gravitation the expressions for the energy, current etc. will have a physical significance only if we define as unity of the gravity charge $\zeta = \sqrt{\gamma} \cdot m$ not the quantity $\zeta_0 = \sqrt{\gamma} \cdot m_0$ belonging to an electron alone, but the gravitational charge $\zeta_1 = \sqrt{\gamma} \cdot Nm_0 = \sqrt{32 \pi \hbar c}$ of an assembly of N electrons such as that of equation (10).

The total charge $\zeta = n \cdot \zeta_1$ moving about a mean orbit of its associated body gives rise to a field, which can be described by a potential four-vector V_i and its complex conjugates; the time component of this $V_0 = n \cdot \zeta_1 / r$ will then agree with Newton's scalar potential.

The conservation laws for the gravitational charge ζ and the potential function V are to be satisfied by an equation of continuity:

$$\operatorname{div} (\zeta v) + \frac{\partial \zeta}{\partial t} = 0, \quad \operatorname{div} V + \frac{1}{c} \frac{\partial V_0}{\partial t} = 0.$$

Through the curl formation of the potential V we may define an anti-symmetric tensor G_{ik} of second order with six components ⁽¹⁾, such that

$$G_{ik} = \frac{\partial V_k}{\partial x_i} - \frac{\partial V_i}{\partial x_k} = (\operatorname{curl} V_i), \quad G_{ik}^* = \frac{\partial V_k^*}{\partial x_i} - \frac{\partial V_i^*}{\partial x_k};$$

G_{ik} is formed by two space vectors Y_x, Y_y, Y_z and U_x, U_y, U_z , together with their complex conjugates.

On the variables Y, U , and V would operate the quantum mechanical momenta

$$p_x = -i\hbar \frac{\partial}{\partial x}, \quad p_y = -i\hbar \frac{\partial}{\partial y}, \quad p_z = -i\hbar \frac{\partial}{\partial z}$$

and the energy operator

$$W = i\hbar \frac{\partial}{\partial t}.$$

Further, we set up the following linear differential equations for the G -field produced by electron assemblies:

$$\left. \begin{aligned} \operatorname{curl} Y + \frac{1}{c} \dot{U} &= 0, & \operatorname{div} Y &= 4\pi \zeta, \\ \operatorname{div} U &= 0, & U &= \operatorname{curl} V, \\ \operatorname{curl} U - \frac{1}{c} \dot{Y} &= \frac{4\pi}{c} \zeta v, & Y + \frac{1}{c} \dot{V} &= -\operatorname{grad} V_0; \end{aligned} \right\} \quad (14)$$

together with the relations for V^*, Y^* , and U^* ; v is the four-velocity of any body of a mass m and a charge ζ .

Each one electron of the system will obey Dirac's ordinary linear wave equation. It can easily be shown that in a plane gravitational wave and for a superposition of G-waves $\text{div } \mathbf{Y} = 0$ is valid.

Equations (14) should induce the second-order wave equations

$$\Delta V = \frac{1}{c^2} \frac{\partial^2 V}{\partial t^2} = -\frac{4\pi}{c} \zeta v, \quad \Delta V_0 - \frac{1}{c^2} \frac{\partial^2 V_0}{\partial t^2} = -4\pi \zeta, \quad (15)$$

the function $V(x, y, z, t)$ being a static solution with central symmetry $1/r$ of this wave equation.

The gravitation current density is to be expressed by a four vector w_k :

$$w_k = \frac{\zeta c i}{\hbar} (G_{ik}^* V^r - G_{ik} V^{r*}), \quad . \quad . \quad . \quad . \quad . \quad (16)$$

provided its time component represents the charge density

$$\tau = -\frac{\zeta i}{\hbar} (G_{ik}^* V^r - G_{ik} V^{r*}); \quad . \quad . \quad . \quad . \quad . \quad (17)$$

ζ is the absolute value of the gravitational charge which belongs to an electronic system.

This enables us to apply to the motion of an electronic assembly in the G-field of another like body the set of equations of the quantum theory.

First, we define a four-vector $K_i = \frac{1}{c} \sum_{k=1}^4 G_{ik} w^k$, which will satisfy the conservation law of energy W and momentum p for the G-field by means of a symmetric tensor T_{ik} , as usual

$$\sum_{k=1}^4 \frac{\partial T_{ik}}{\partial x^k} = K_i = \frac{1}{c} \sum_{k=1}^4 G_{ik} w^k.$$

The stress-energy tensor T_{ik} shall then have the following components:

$$\left. \begin{aligned} T_{xx} &= \frac{1}{4\pi} [Y_x^2 + U_x^2 - \frac{1}{2}(Y^2 + U^2)], \text{ and similarly for } T_{yy}, T_{zz}, \\ T_{xy} &= \frac{1}{4\pi} (Y_x Y_y + U_x U_y) = T_{yx}, \\ T_{00} &= c(\frac{1}{2} G_{ik}^* G_{ik} + Y^* Y + V_i^* V_i + V_0^* V_0); \end{aligned} \right\} \quad . \quad . \quad (18)$$

among these the 00-component of T_{ik} is the energy density of the field. A Lagrangian function is required for establishing the action principle; let this be given by

$$L = \frac{1}{4\pi} (Y^* Y - U^* U) - (V_i^* V_i + V_0^* V_0), \quad . \quad . \quad . \quad (19)$$

whereas the fundamental function of the field will be the Hamiltonian \mathbf{H} , namely, the space integral of \mathbf{T}_{00} :

$$\begin{aligned}\mathbf{H} &= \int \mathbf{T}_{00} dv = \left[P \frac{\partial V}{\partial t} + P_0 \frac{\partial V_0}{\partial t} + \text{compl. conj.} - \mathbf{L} \right] \\ &= c \int dv \left[\mathbf{Y}^* \mathbf{Y} + \frac{1}{2} \left(\frac{\partial V_i^*}{\partial x_k} - \frac{\partial V_k^*}{\partial x_i} \right) \left(\frac{\partial V_i}{\partial x_k} - \frac{\partial V_k}{\partial x_i} \right) + (V_i^* V_i + V_0^* V_0) \right]. \quad (20)\end{aligned}$$

The functions \mathbf{L} and \mathbf{H} correspond to the Lagrangian one which first had been proposed by Proca ⁽⁵⁾ in his relativistic field theory of the electron.

We might eliminate V_0 and V_0^* from (20) by imposing the conditions

$$V_0 = -\text{div } \mathbf{Y}^*, \quad V_0^* = -\text{div } \mathbf{Y}; \quad \dots \dots \dots (21)$$

hence we obtain

$$\mathbf{H} = c \int dv [\mathbf{Y}^* \mathbf{Y} + \text{curl } \mathbf{V}^* \text{ curl } \mathbf{V} + V_i^* V_i + \text{div } \mathbf{Y}^* \text{div } \mathbf{Y}]. \quad (22)$$

This function leads up to the Hamilton equations which shall govern the motion of a body in the gravitational field of a second body :

$$\begin{aligned}\dot{q} = \partial x / \partial t &= \partial \mathbf{H} / \partial p_x, \text{ i. e., } \frac{\partial V_x}{\partial t} = \frac{1}{i\hbar} (V_x \mathbf{H} - \mathbf{H} V_x) \text{ etc.} \\ &= Y_x - \frac{\partial}{\partial x_i} \text{div } \mathbf{Y} \text{ etc.} \quad \dots \dots \dots (23)\end{aligned}$$

$$\begin{aligned}\dot{p} = \partial p_x / \partial t &= -\partial \mathbf{H} / \partial x, \text{ i. e., } \frac{\partial Y_x}{\partial t} = \frac{1}{i\hbar} (Y_x \mathbf{H} - \mathbf{H} Y_x) \text{ etc.} \\ &= V_x + \frac{\partial}{\partial x_i} \text{curl}_x \mathbf{V} - \frac{\partial}{\partial x_i} \text{curl}_y \mathbf{V} \text{ etc.} \quad (24)\end{aligned}$$

6. Quantization of the G-field Equations.

We thus choose as canonical coordinates to be the (negative) vector potential V_i or V_i^* , as their conjugate momenta p the field strengths $\mathbf{Y}/4\pi c$ or $\mathbf{Y}^*/4\pi c$, and we now proceed to the quantization of the preceding equations.

Since the propagation of gravitational actions, we assume, takes place through the intermediary of a quantum which we allow to have an integer spin, in close analogy with the propagation of electromagnetic perturbations by light quanta, the field variables V_i , V_i^* and Y_k , Y_k^* must therefore obey commutability relations for Einstein-Bose Statistics.

To the same conclusion leads also an application of the quantum mechanical formalism to the general relativity theory by L. Rosenfeld ⁽⁶⁾.

Those commutation rules are as follows :—

$$\left. \begin{aligned} [V_i(x), Y^*(x') - V_i^*(x'), Y_k(x)] &= 4\pi i\hbar c \delta(x-x') \delta_{ik}, \\ [Y_k(x), U^*(x') - U_k^*(x'), Y_i^*(x)] &= i\hbar \frac{\partial}{\partial x'} \delta(x-x') \end{aligned} \right\} \quad . \quad . \quad (25)$$

whereas all the other Poisson-Brackett's relations commute between themselves.

The commutability laws (25) must be relativistically invariant, and such a proof is to be established if we introduce an infinitesimal anti-symmetric tensor $\mu_{ik} = -\mu_{ki}$ of rank two; the six components of this tensor should be integrals of motion and will correspond to the group of six Lorentz transformations, as in the quantum electrodynamics⁽⁷⁾,

The field equations (14) can be integrated by approximations in which we resolve the field strengths at a given time into Fourier series. Let us impose as boundary condition the ordinary periodicity of the wave functions in respect to any cube of side L ; the wave-length $1/K_r$ of the oscillator will then be a whole multiple of $L/2\pi$.

For a complete orthogonal system we choose that of the coordinates, namely, the vector potential components

$$V_i = c \left(\frac{8}{L} \right)^{\frac{1}{2}} q_i^r \cos \frac{\pi}{L} k_1 x_1 \cdot \sin \frac{\pi}{L} k_2 x_2 \cdot \sin \frac{\pi}{L} k_3 x_3, \quad i=1, 2, 3, 4, \quad . \quad (26)$$

and similarly for the canonical conjugate momenta p .

The Hamiltonian (20) of the gravitational field owing to electron assemblies would become

$$\mathbf{H} = \hbar c k \sum_k \sum_{j=-1}^1 [q_j^*(k) q_i(k) + p_j^*(k) p_j(k)], \quad . \quad . \quad . \quad (27)$$

and similarly the total momentum

$$\mathbf{P}_r = -i\hbar k \sum_k \sum_{j=-1}^1 [p_j(k) q_j(k) - q_j^*(k) p_j^*(k)]. \quad . \quad . \quad . \quad (28)$$

The canonical conjugate operators p and q will, of course, obey commutability laws as those of (25). The total gravity current (16) reveals itself as

$$\frac{1}{c} \bar{\zeta} v_r = \int w_r dv = \frac{\zeta c i}{\hbar} \sum_k \sum_{j=-1}^1 [p_j(k) q_j(k) - p_j^*(k) q_j^*(k)], \quad . \quad . \quad (29)$$

whereas the total gravity charge (17) will be stated as

$$\bar{\zeta} = \frac{1}{c} \int \tau dv = -\frac{\zeta i}{\hbar} \sum_k \sum_{j=-1}^1 [p_j(k) q_j(k) - p_j^*(k) q_j^*(k)]. \quad . \quad . \quad . \quad (30)$$

If we wish to explain the facts both of an integer spin $1 \hbar$ having been allowed to the gravitation quanta as well as of the polarization of the

gravitational waves, we have to separate each one space quantity of the field into a part polarized transversely and a part longitudinally. In particular, the potential space vector will accordingly be broken up, thus

$$V_r = \psi_r^{tr} + \psi_r^{*long}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

and also the field strength

$$Y_r^* = \Pi_r - \pi_r^*. \quad . \quad . \quad . \quad . \quad . \quad . \quad (32)$$

In a similar way (25) to be represented the Hamiltonian (22) and the commutation rules are. The suffix j in the previous equations (27) or (29) will take three values 0, ± 1 of the spin, the first zero belonging to longitudinal quanta, the other two to transversal G-quanta with a perpendicular polarization.

7. The Gravitational Field due to Atomic Nuclei Systems.

We now consider the question of the presence of nuclear systems, that is, of protons or neutrons inside the gravitational field of electron assemblies (§5) and their mutual interaction. We shall need to complete the equation (14) by a new antisymmetric tensor M_{ik} and by a four-vector χ_i which may characterize each one heavy particle of the system. Neutrons and protons will again conform to Dirac's linear equation involving their wave functions Φ_N and Φ_p .

Each component of the vector χ_i should have eight components, and the four components of χ_i are defined as

$$\chi_r = (\Phi_N^* \alpha_i Q^* \Phi_p), \quad \chi_0 = (-\Phi_N^* Q^* \Phi_p), \quad . \quad . \quad . \quad (33)$$

α_i being the spin vector and Q^* the operator of the proton or neutron state.

On the other hand, the tensor M_{ik} is formed by two space vectors :

$$M_r = (\Phi_N \rho_2 \sigma_r Q^* \Phi_p), \quad R_r = (\Phi_N^* \rho_3 \sigma_r Q^* \Phi_p), \quad . \quad . \quad . \quad (34)$$

where σ_r , ρ_2 , and ρ_3 are the well-known Dirac matrices.

Therefore, we can set up the equations of the gravitational field produced by assemblies of heavy particles and electrons in the form :

$$\left. \begin{aligned} \text{curl } U - \frac{1}{c} \dot{Y} &= 4\pi \zeta_1 \chi_r, & Y + \frac{1}{c} \dot{V}_i + \text{grad } V_0 &= 4\pi \zeta_2 M_r \\ \text{div } Y &= 4\pi \zeta_1 \chi_0, & U - \text{curl } V &= 4\pi \zeta_2 \bar{R}_r; \end{aligned} \right\} \quad (35)$$

here ζ_1 and ζ_2 play the part of a charge for the longitudinal or transversal waves of gravitation.

These field equations must be derived from the relativistically invariant Lagrangian function (19), completed by M_{ik} and χ_i :

$$L = \frac{1}{4\pi} (Y^*Y - U^*U) - (V_i^*V_i + V_0^*V_0) + \zeta_1(V_r^*\chi_r - V_0^*\chi_0 + V_r\chi_r^* - V_0\chi_0^*) + \zeta_2(\frac{1}{2}M_{ik}G_{ik}^* + \frac{1}{2}M_{ik}^*G_{ik}). \quad (36)$$

The Hamiltonian function of the gravitational field of nuclear and electronic matter will be given by the sum :

$$H = H_G + H_H + H'_{1,2}. \quad (37)$$

H_G is the function (20) of electron systems, while H_H represents the function for neutron and proton systems :

$$H_H = \Psi^* \left[c\alpha_i p_i + \beta \left(\frac{1+\theta_3}{2} \chi_N c^2 + \frac{1-\theta_3}{2} \chi_p c^3 \right) \right] \Psi. \quad (38)$$

α_i and β 's are the Dirac spin matrices, whereas θ_3 means the isotopic spin matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which for $\theta_3=1$ gives the neutron state and for $\theta_3=-1$ the proton state.

The function $H'_{1,2}$ of gravitational interaction will involve terms of transversal and longitudinal gravitation-waves in the first- and second-order powers of the charges ζ_1, ζ_2 ; this will be made if we leave the right-hand sides of (35) to operate on H_G , in which (20) we eliminate V_0^* and V_0 by the definitions

$$V_0^* = -\text{div } Y + \zeta_1 \chi_0^*, \quad V_0 = -\text{div } Y^* + \zeta_1 \chi_0. \quad (21')$$

The interaction function is thus found to be

$$H'_{1,2} = \zeta_1 V^* \chi_r + \zeta_1 V \chi_r^* + \zeta_2 Y M_r + \zeta_2 Y^* M_r^* + \zeta_2 \text{curl } V R_r^* + \zeta_2 \text{curl } V^* R_r + \zeta_1 \text{div } Y \chi_0 + \zeta_1 \text{div } Y^* \chi_0^* + \zeta_1^2 \chi_r^* \chi_r + \zeta_1^2 \chi_0^* \chi_0 + \zeta_2^2 M_r^* M_r + \zeta_2^2 R_r^* R_r. \quad (39)$$

The interaction between two material systems takes place by simultaneous emission or absorption of longitudinal and transversal gravitation quanta ; the Newton law $W = -\gamma \cdot m^2/r = -(\zeta_1^2 + \zeta_2^2)/r$ can be deduced from the total Hamiltonian (37) in the perturbation energy of second order, as the Coulomb law may be derived from the quantized Maxwell-Lorentz equations⁽⁸⁾. This will be given on another occasion.

8. The Cosmological Equations.

For distances $r \rightarrow \infty$ Newton's law should take the exponential form (11) :

$$\zeta^2 \frac{e^{-\kappa r}}{r}, \text{ with } \kappa = m_g c / \hbar = 1/2\pi R_0.$$

m_g signifies the very small rest mass of the G-quantum, while R_0 is the large length (13).

The second-order wave equations (15) will be followed by

$$\Delta V - \frac{1}{c^2} \frac{\partial^2 V}{\partial t^2} - \kappa^2 V = \frac{4\pi}{c} \zeta v, \quad \Delta V_0 - \frac{1}{c^2} \frac{\partial^2 V_0}{\partial t^2} - \kappa^2 V_0 = \frac{4\pi}{c} \zeta, \quad (40)$$

and similarly for the linear equations (14) of the gravitational field owing to electron assemblies.

We may also take into account the G-field of systems of heavy particles separated from one another by a large length of the order of magnitude of the cosmical radius R_0 . This can be done by setting up the equations (35) in the cosmological form :

$$\left. \begin{aligned} \kappa V_r - \frac{1}{c} \dot{Y}_r + \text{curl } U &= 4\pi \zeta_1 \chi_r, & \kappa Y + \frac{1}{c} \dot{V}_i + \text{grad } V_0 &= 4\pi \zeta_2 M_r, \\ \kappa V_0 + \text{div } Y &= 4\pi \zeta_1 \chi_0, & \kappa U - \text{curl } V &= 4\pi \zeta_2 R_r. \end{aligned} \right\} \quad (41)$$

The Lagrangian (36) and also the Hamiltonian (22) or (39) will here include the terms $\frac{\zeta_1}{\kappa}$ and $\frac{\zeta_2}{\kappa}$ instead of ζ_1 or ζ_2 , whereas in the latter fractions as $\frac{\zeta_1^2}{\kappa^2}$ or $\frac{\zeta_2^2}{\kappa^2}$ should replace the squares ζ_1^2 or ζ_2^2 .

The commutability laws (25) are left unchanged, but the equations of motion (23) and (24) will become

$$\frac{\partial V_i}{\partial t} = \kappa Y - \frac{\partial}{\partial x_i} Y, \quad \frac{\partial Y^*}{\partial t} = V_i + \frac{1}{\kappa} \frac{\partial}{\partial x_i} \text{curl}_r V_i - \frac{1}{\kappa} \frac{\partial}{\partial x_i} \text{curl}_s V_i. \quad (42)$$

The square of the matrix elements of the interaction H' gives us the probability of emission or absorption of a G-quantum in the form of a plane wave $a \exp i(\kappa r)$ by a system of particles ; these matrix elements may then be written as

$$-Q^* \frac{\zeta_1}{\kappa} \sqrt{\left(\frac{2\pi\hbar^2}{m_g}\right)} e^{-i(\kappa r)} \cdot \kappa, \quad -Q^* \frac{\zeta_2}{\kappa} \sqrt{\left(\frac{2\pi\hbar^2}{m_g}\right)} e^{-i(\kappa r)} (\sigma[a\kappa])$$

for a longitudinal and a transversal G-quantum respectively.

The complex conjugates to these expressions hold for the case of absorption. The inter-reaction operator H can also be expressed by a longitudinal and a transversal part, for instance

$$H_{\text{long}} = -2 \frac{\zeta_1^2}{\kappa^2} 2\pi\hbar Q^* \Sigma \frac{\kappa^2 e^{i\kappa r}}{m_g} = +\zeta_1^2 Q^* \frac{e^{-\kappa r}}{r}.$$

Up to the present we have developed a straightforward analogy between cosmical gravitation and nuclear physics : the emission of gravitational quanta by systems of electrons and heavy particles would correspond to the emission of γ -rays by atomic nuclei or to the emission of Bose particles (heavy electrons) by neutrons and protons.

9. *The Total Gravitational and Electromagnetic Field.*

We wish again to regard the presence of the electromagnetic field, provided with a four-potential⁽¹⁾ vector $A_i=A_a$, A_0 inside the field of gravitation.

The gravitational tensor G_{ik} will then be followed, in both ordinary and cosmological cases, by

$$\left(\frac{\partial}{\partial x^a} - \frac{ie}{\hbar c} A_a\right) V_\beta - \left(\frac{\partial}{\partial x^\beta} - \frac{ie}{\hbar c} A_a\right) V_a = G_{ik} \text{ or } = \kappa G_{ik}, \quad (43)$$

$$\left(\frac{\partial}{\partial x_a} - \frac{ie}{\hbar c} A_a\right) G_{ik} = V^\beta \text{ or } = \kappa V^\beta \quad (44)$$

grad and $\frac{1}{c} \frac{\partial}{\partial t}$, which operated on the G-field variables Y , U , Y_r , or V_0 should be replaced by grad

$$-\frac{ie}{\hbar c} A_a \text{ and } \frac{1}{c} \frac{\partial}{\partial t} + \frac{ie}{\hbar c} A_0.$$

The field equations which replace (14) will obviously be written

$$\left. \begin{aligned} \left(\text{grad} + \frac{ie}{\hbar c} A_a\right) U_r - \left(\frac{1}{c} \frac{\partial}{\partial t} - \frac{ie}{\hbar c} A_0\right) Y_r - \frac{4\pi}{c} (\zeta + \rho) &= -V_r \text{ or } = -\kappa V_r, \\ \left(\text{grad} - \frac{ie}{\hbar c} A_a\right) Y_r - \frac{4\pi}{c} (\zeta + \rho) &= -V_0 \text{ or } = -\kappa V_0, \\ \left(\text{grad} - \frac{ie}{\hbar c} A_a\right) V_r &= U_r \text{ or } = \kappa U_r, \\ \left(\text{grad} - \frac{ie}{\hbar c} A_a\right) V_0 + \left(\frac{1}{c} \frac{\partial}{\partial t} + \frac{ie}{\hbar c} A_0\right) V_r &= -Y_r \text{ or } = -\kappa Y_r, \end{aligned} \right\} \quad (45)$$

and their complex conjugates. ρ means the electric charge density.

It can be proved that these field equations should be derived by variation of the gravitational potentials V_i , V_i^* from a Lagrangian of the complete field as :

$$\bar{\mathbf{L}} = \mathbf{L}_G + \mathbf{L}_E,$$

whereas \mathbf{L}_G is the ordinary or cosmological form of the function (19) for the gravitational field due to electron assemblies ; \mathbf{L}_E is the classical expression

$$\mathbf{L}_E = \frac{1}{8\pi} (\mathbf{E}^2 - \mathbf{H}^2)$$

This ensures that the same action principle, if we vary the electro-

dynamic potentials A_i , leads up to the Maxwell's linear equations with a ρ -charge density :

$$\begin{aligned} E_r &= -\frac{1}{c} \frac{\partial A_r}{\partial t} - \text{grad } A_0, & H &= \text{curl } A, \\ \text{curl } H &= \frac{1}{c} \frac{\partial E_r}{\partial t} + \frac{4\pi}{c} I_r, & \text{div } E &= 4\pi \rho. \end{aligned}$$

The complete charge density carried by electron and proton or neutron systems when interacting between themselves will be found as the sum both of electric and gravitational charges, the latter expressed in quantum unities :

$$\begin{aligned} \rho + \tau &= \frac{i(e+\zeta)}{\hbar c} \left\{ V^* \frac{1}{c} \frac{\partial V}{\partial t} - \frac{1}{c} \frac{\partial V^*}{\partial t} V - \frac{ie}{\hbar c} A_0 V^* V_0 - \frac{ie}{\hbar c} (AV) V_0^* \right. \\ &\quad \left. - V^* \text{grad } V_0 + \text{grad } V_0^* V + \frac{ie}{\hbar c} (V^* A) V_0 + \frac{ie}{\hbar c} V_0^* (AV) \right\}, \quad (46) \end{aligned}$$

whereas the complete current density should be

$$\begin{aligned} I_x + w_x &= -\frac{i(e+\zeta)}{\hbar} \left\{ V^* \frac{\partial V}{\partial x} - \frac{\partial V^*}{\partial x} V - \frac{ie}{\hbar c} A_x V^* V_0 - \frac{ie}{\hbar c} A_x V V_0^* \right. \\ &\quad \left. - (V^* \text{grad}) V_x + (V \text{grad}) V_x^* + \frac{ie}{\hbar c} (V^* A) V_x + \frac{ie}{\hbar c} V_x^* (AV) \right\} \quad (47) \end{aligned}$$

and so on.

For a convenient Hamiltonian function of the gravitational and electromagnetic field one gets the sum

$$\bar{H} = H_{G,A} + H_E. \quad (48)$$

$H_{G,A}$, which involves interaction terms like $\frac{ie}{\hbar} Y A_0 V$, would be given by means of (20) and (45) :

$$\begin{aligned} H_{G,A} &= (\kappa) c \int dv \left[Y^* Y - \frac{ie}{\hbar} Y A_0 V - \left\{ \left(\text{grad} - \frac{ie}{\hbar c} A_a \right) Y_r \right\} \right. \\ &\quad \times \left\{ \text{grad} + \frac{ie}{\hbar c} A_0 \right\} Y_r^* - \frac{ie}{\hbar} Y^* A_0 V^* - \left(\frac{1}{\kappa^2} \right) \left. \right\} \\ &\quad \times \left\{ \left(\text{grad} + \frac{ie}{\hbar c} A_a \right) V^* \right\} \left\{ \left(\text{grad} - \frac{ie}{\hbar c} A_a \right) V + V_i^* V_i \right\} \Big], \quad (49) \end{aligned}$$

in which the factors (κ) and $\left(\frac{1}{\kappa^2}\right)$ may emphasize the cosmological case.

H_E in (48) is the classical Hamiltonian $\frac{1}{8\pi} (E^2 + H^2)$.

The variables V_i and Y_r of the gravitational field will again satisfy the commutation rules as (25); analogously, the motion of the system

of electrically charged particles inside the field of gravitation is to be guided by the Hamiltonian equations (37).

In dealing with electric charges it is convenient to exhibit the theory of the second quantification in the electromagnetic part of the total field, and thus to interpret the Maxwell field in terms of numbers of electric charges. Let us endeavour to establish instead of the canonical conjugate variables p , p^* , and q , q^* the new variables a , a^* and b , b^* through the Pauli and Weisskopf's transformation⁽⁹⁾ :

$$p_j(k) = \left(\frac{\hbar ck}{2}\right)^{\frac{1}{2}} [a_j^*(k) + b_j(k)], \quad q_j(k) = i \left(\frac{\hbar}{2k}\right)^{\frac{1}{2}} [a_j^*(k) - b_j(k)].$$

Those definitions of the new variables allow them to satisfy the simple commutability relations

$$[a_j(k), a_j^*(k')] = [b_j(k), b_j^*(k)] = \delta_{jj}, \quad \delta(k, k'),$$

all the other combinations cancelling.

The Hamiltonian is merely written in the form

$$\mathbf{H} = \sum_k \sum_{j=-1}^1 \hbar ck [n_j^+(k) + n_j^-(k) + 1], \quad . \quad . \quad . \quad . \quad (50)$$

from which and from the commutability relations we draw the conclusion that the operators

$$n_j^+(k) = a_j^*(k) a_j(k), \quad n_j^-(k) = b_j^*(k) b_j(k)$$

have positive eigenvalues 0, 1, 2 . . . Since the total electric charge is here found to be

$$\bar{e} = \frac{1}{c} \int \rho dv = e \sum_k \sum_{j=-1}^1 [a_j^*(k) a_j(k) - b_j^*(k) b_j(k)],$$

thus the operators n^+ and n^- express the number of electric particles according to whether their charge is positive or negative.

References.

- (1) G. H. Livens and I. E. Viney, *Phil. Mag.* xiv. pp. 243-254 (1932) ; G. H. Livens, I. E. Viney, and G. G. Leybourne, *loc. cit.* xv. pp. 33-48 (1933) ; G. H. Livens and H. F. Willis, *loc. cit.* xv. p. 142 (1933).
- (2) G. Gamow, 'Radioactivity' (Oxford), second ed. p. 145 (1937).
- (3) H. Flint, *Proc. Roy. Soc. Lond., A*, clx. p. 45 (1937) ; A. March, *Z. Physik*, cviii. p. 128 (1937), etc.
- (4) R. Millikan, *Ann. der Phys.* xxxii. pp. 520 & 34 (1938).
- (5) Al. Proca, *J. Phys. Radium*, vii. p. 347 (1936).
- (6) L. Rosenfeld, *Ann. der Phys.* v. p. 150 (1930).
- (7) M. H. L. Pryce, *Proc. Roy. Soc. Lond., A*, clx. p. 371 (1937).
- (8) E. Fermi, *Rev. Mod. Phys.* iv. p. 131 (1932).
- (9) W. Pauli and V. Weisskopf, *Helv. Phys. Acta*, vii. p. 709 (1934).

Institute of Theoretical Physics
of the University of Athens.

*XLIII. Non-linear Theory of Elasticity and the Linearized Case
for a Body under Initial Stress.*

By Prof. M. A. BIOT *.

[Received December 23, 1938.]

Introduction.

IT is well known that the classical Theory of Elasticity is restricted to small deformations and rotations and that this is the underlying reason for its linear character. Attempts have been made ⁽¹⁾ to establish a theory for finite deformations starting from the general mathematical viewpoint of tensor theory.

It was thought that a non-linear theory including terms of the first and second order only would yield the essential features due to large deformations which are not explained by a linear theory. Such features are exhibited, for instance, in the flexure of a thin shell.

Our method does not require an explicit formulation of the stress-strain relation, which is a physical problem. The essential idea which led us to our equations was to refer the stress condition to a local system of axis rotating with the material at that point, and to investigate equilibrium conditions for these stress components. These equations will contain explicitly only those second-order terms which are of kinematic origin. This development is made in section 2, while section 1 deals with a more accurate definition of stress and strain. This first section introduces a new definition of the strain components, which are here linearly related to the actual change of distance between two neighbouring points in the material.

These strain components are very important in establishing a correct expression for the potential energy and deriving the equilibrium equations by the variational method. This derivation is made in section 3, and the same equations are found as in the previous method. However, consideration of the strain energy leads to a new interesting viewpoint, as it introduces naturally two forms of representation of the stress—one in which the stresses are referred to the actual areas after deformation, and these are the stresses adopted in the previous section 2, the other in which the stresses are referred to the areas before deformation. Relations are found between these two types of stress components, and equations of equilibrium derived for both.

* Communicated by the Author.

In section 4 we carry our attention to a linear theory of elasticity for bodies under high initial stress. This covers a wide range of problems, from elastic stability and buckling to the propagation of elastic waves inside the earth and the meaning of the elastic coefficients for a body under high initial stress. The special problem of elastic stability has been the object of many previous researches. R. V. Southwell⁽²⁾ considers a uniform initial stress, and chooses axes along the principal directions; E. Trefftz⁽³⁾ uses the variational method, but he chooses an expression for the potential energy which is correct only if the rotation is large with respect to the strain; C. B. Biezeno and H. Hencky⁽⁴⁾ have developed a general theory.

The reader will be aware that methods similar to those in the previous sections may be used to establish the linear theory of elasticity of a body under high initial stress. In fact the equations can be derived from the previous theory by linearizing with respect to small increments of stress and the small components of strain and rotation. Those of our equations using the stress components referred to the initial areas are found to coincide with those derived by C. B. Biezeno and H. Hencky⁽⁴⁾ by an entirely different method.

The last section (5) deals with the special case when we wish to introduce from the start the assumption that the rotation is large with respect to the strain. Special precautions have to be taken in introducing this assumption, because the strain and the rotation are not independent and satisfy the identities (2.6). Equations are derived which are a generalization of the result obtained by the author⁽⁵⁾ in a previous paper.

1. *Strain and Stress.*

The original coordinates x, y, z of a point attached to the material become

$$\begin{aligned}\xi &= x + u, \\ \eta &= y + v, \\ \zeta &= z + w\end{aligned}$$

after the deformation. The infinitesimal region surrounding this point undergoes a homogeneous deformation defined by the linear transformation of dx, dy, dz into $d\xi, d\eta, d\zeta$,

$$\left. \begin{aligned}d\xi &= \left(1 + \frac{\partial u}{\partial x}\right) dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz, \\ d\eta &= \frac{\partial v}{\partial x} dx + \left(1 + \frac{\partial v}{\partial y}\right) dy + \frac{\partial v}{\partial z} dz, \\ d\zeta &= \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \left(1 + \frac{\partial w}{\partial z}\right) dz,\end{aligned}\right\} \dots \dots (1.1)$$

$$\text{or} \quad \left. \begin{aligned} d\xi &= (1+e_{xx}) dx + (e_{xy}-\omega_z) dy + (e_{xz}+\omega_y) dz, \\ d\eta &= (e_{xy}+\omega_z) dx + (1+e_{yy}) dy + (e_{yz}-\omega_x) dz, \\ d\zeta &= (e_{xz}-\omega_y) dx + (e_{yz}+\omega_x) dy + (1+e_{zz}) dz, \end{aligned} \right\} \quad . \quad . \quad . \quad (1.2)$$

$$\text{with} \quad \left. \begin{aligned} e_{xx} &= \frac{\partial u}{\partial x}, & e_{yz} &= \frac{1}{2} \left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \right), & \omega_z &= \frac{1}{2} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right), \\ e_{yy} &= \frac{\partial v}{\partial y}, & e_{zx} &= \frac{1}{2} \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right), & \omega_y &= \frac{1}{2} \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right), \\ e_{zz} &= \frac{\partial w}{\partial z}, & e_{xy} &= \frac{1}{2} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right), & \omega_x &= \frac{1}{2} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right). \end{aligned} \right\} \quad . \quad . \quad (1.3)$$

The length element after the deformation is

$$ds^2 = (1+2g_{xx}) dx^2 + (1+2g_{yy}) dy^2 + (1+2g_{zz}) dz^2 \\ + 4g_{yz} dy dz + 4g_{zx} dz dx + 4g_{xy} dx dy, \quad . \quad . \quad . \quad (1.4)$$

$$\text{with} \quad \left. \begin{aligned} g_{xx} &= e_{xx} + \frac{1}{2}e_{xx}^2 + \frac{1}{2}(e_{xy}+\omega_z)^2 + \frac{1}{2}(e_{xz}-\omega_y)^2, \\ g_{yy} &= e_{yy} + \frac{1}{2}e_{yy}^2 + \frac{1}{2}(e_{yz}+\omega_x)^2 + \frac{1}{2}(e_{xy}-\omega_z)^2, \\ g_{zz} &= e_{zz} + \frac{1}{2}e_{zz}^2 + \frac{1}{2}(e_{zx}+\omega_y)^2 + \frac{1}{2}(e_{yz}-\omega_x)^2, \\ g_{yz} &= e_{yz} + \frac{1}{2}(e_{xy}-\omega_z)(e_{zx}+\omega_y) + \frac{1}{2}e_{yy}(e_{yz}-\omega_x) + \frac{1}{2}e_{zz}(e_{yz}+\omega_x), \\ g_{zx} &= e_{zx} + \frac{1}{2}e_{xx}(e_{zx}+\omega_y) + \frac{1}{2}(e_{yz}-\omega_x)(e_{xy}+\omega_z) + \frac{1}{2}e_{zz}(e_{zx}-\omega_y), \\ g_{xy} &= e_{xy} + \frac{1}{2}e_{xx}(e_{xy}-\omega_z) + \frac{1}{2}e_{yy}(e_{xy}+\omega_z) + \frac{1}{2}(e_{zx}-\omega_y)(e_{yz}+\omega_x). \end{aligned} \right\} \quad (1.5)$$

The transformation (1.2) contains the *nine* independent coefficients (1.3), while the change of length ds^2 depends only on the six coefficients g . There are therefore three degrees of freedom, leaving unchanged the length element ds^2 and corresponding to the rigid body rotation contained in the general linear transformation (1.2).

One of the first questions arising in the theory of elasticity is to distinguish what part in the general transformation (1.2) is to be considered as a pure strain and what part as a pure rotation. It is well known that when the quantities (1.3) are all small of the first order, and when we neglect quantities of higher order the pure deformation is represented by the coefficients e and the rotation by the vector $\omega_x \omega_y \omega_z$. However, this is only a first approximation. In the following we are concerned with finding the finite pure deformation contained in the transformation (1.2) and developing an expression for the strain components containing both first and second order terms. We consider therefore the following linear transformation :

$$\left. \begin{aligned} d\xi_1 &= (1+\epsilon_{11}) dx + \epsilon_{12} dy + \epsilon_{31} dz, \\ d\xi_2 &= \epsilon_{12} dx + (1+\epsilon_{22}) dy + \epsilon_{23} dz, \\ d\xi_3 &= \epsilon_{31} dx + \epsilon_{23} dy + (1+\epsilon_{33}) dz, \end{aligned} \right\} \quad . \quad . \quad . \quad (1.6)$$

with symmetric coefficients. Such a transformation leaves unchanged the orientation of three rectangular directions which are the principal directions of strain ⁽⁶⁾; in other words, the transformation (1.6) is equivalent to elongations along three fixed rectangular directions. It is therefore quite natural to call such a transformation a "pure deformation." We may choose as finite strain components of the deformation the symmetric coefficients

$$\left. \begin{array}{ccc} \epsilon_{11}, & \epsilon_{12}, & \epsilon_{31}, \\ \epsilon_{12}, & \epsilon_{22}, & \epsilon_{23}, \\ \epsilon_{31}, & \epsilon_{23}, & \epsilon_{33}. \end{array} \right\} \quad . \quad . \quad . \quad . \quad . \quad (1.7)$$

The length of an element after the transformation (1.6) is

$$d\sigma^2 = (1 + 2\gamma_{11}) dx^2 + (1 + 2\gamma_{22}) dy^2 + (1 + 2\gamma_{33}) dz^2 \\ + 4\gamma_{23} dydz + 4\gamma_{31} dzdx + 4\gamma_{12} dxdy, \quad . \quad . \quad . \quad (1.8)$$

with

$$\left. \begin{array}{l} \gamma_{11} = \epsilon_{11} + \frac{1}{2}(\epsilon_{11}^2 + \epsilon_{12}^2 + \epsilon_{31}^2), \\ \gamma_{22} = \epsilon_{22} + \frac{1}{2}(\epsilon_{22}^2 + \epsilon_{12}^2 + \epsilon_{23}^2), \\ \gamma_{33} = \epsilon_{33} + \frac{1}{2}(\epsilon_{33}^2 + \epsilon_{31}^2 + \epsilon_{23}^2), \\ \gamma_{23} = \epsilon_{23} + \frac{1}{2}(\epsilon_{12}\epsilon_{31} + \epsilon_{22}\epsilon_{23} + \epsilon_{33}\epsilon_{23}), \\ \gamma_{31} = \epsilon_{31} + \frac{1}{2}(\epsilon_{11}\epsilon_{31} + \epsilon_{23}\epsilon_{12} + \epsilon_{33}\epsilon_{31}), \\ \gamma_{12} = \epsilon_{12} + \frac{1}{2}(\epsilon_{11}\epsilon_{12} + \epsilon_{22}\epsilon_{12} + \epsilon_{31}\epsilon_{23}). \end{array} \right\} \quad . \quad . \quad . \quad . \quad (1.9)$$

Now the pure deformation (1.6) can be made to represent exactly the same state of strain as that produced by the transformation (1.2), provided the length elements ds^2 and $d\sigma^2$ after deformation are identical. This condition is expressed analytically by the six equations

$$\left. \begin{array}{ll} g_{xx} = \gamma_{11}, & g_{yz} = \gamma_{23}, \\ g_{yy} = \gamma_{22}, & g_{zx} = \gamma_{31}, \\ g_{zz} = \gamma_{33}, & g_{xy} = \gamma_{12}. \end{array} \right\} \quad . \quad . \quad . \quad . \quad (1.10)$$

These equations determine the six strain components (1.7) as functions of the nine quantities (1.3).

The transformations (1.2) and (1.6) thus related represent the same state of strain, and can only differ by a rigid body rotation. The rigid body rotation that we must add to the transformation (1.6) in order to obtain the transformation (1.2) will be called the "local rotation" of the material.

The strain components (1.7) have the advantage that they are linearly related to the actual changes of length in the material, while the classical components (1.5) are linearly related to the change of the square of the

length. On the other hand, they have the disadvantage that they cannot be expressed rationally by means of the nine quantities (1.3). However, this disadvantage vanishes in one important case—that is, when we assume the *nine quantities* (1.3) to be *small of the first order* and when we consider only the first- and second-order terms in the expression of the strain tensor (1.7) as a function of the nine quantities (1.3). In this case the result is obtained immediately as follows.

We notice from equations (1.10) that e_{xx} and e_{11} differ only by a second-order quantity, the same for e_{xy} and e_{12} , etc., so that we may write with an error of the third order :

$$\left. \begin{aligned} e_{xx}^2 + e_{xy}^2 + e_{zx}^2 &= \epsilon_{11}^2 + \epsilon_{12}^2 + \epsilon_{13}^2, \\ e_{yy}^2 + e_{xy}^2 + e_{yz}^2 &= \epsilon_{22}^2 + \epsilon_{12}^2 + \epsilon_{23}^2, \\ e_{zz}^2 + e_{zx}^2 + e_{yz}^2 &= \epsilon_{33}^2 + \epsilon_{31}^2 + \epsilon_{23}^2, \\ e_{xy}e_{xx} + e_{yy}e_{yz} + e_{zz}e_{yz} &= \epsilon_{12}\epsilon_{31} + \epsilon_{22}\epsilon_{23} + \epsilon_{33}\epsilon_{23}, \\ e_{xx}e_{zx} + e_{yz}e_{xy} + e_{zz}e_{xx} &= \epsilon_{11}\epsilon_{31} + \epsilon_{23}\epsilon_{12} + \epsilon_{33}\epsilon_{31}, \\ e_{xx}e_{xy} + e_{yy}e_{xy} + e_{zx}e_{yz} &= \epsilon_{11}\epsilon_{12} + \epsilon_{22}\epsilon_{12} + \epsilon_{31}\epsilon_{23}. \end{aligned} \right\} \quad . \quad . \quad (1.11)$$

Introducing the approximate identities (1.11) into equations (1.10) we find the following expressions for the strain tensor with an error of the third order :

$$\left. \begin{aligned} \epsilon_{11} &= e_{xx} + e_{xy}\omega_z - e_{zx}\omega_y + \frac{1}{2}(\omega_z^2 + \omega_y^2), \\ \epsilon_{22} &= e_{yy} + e_{yz}\omega_x - e_{xy}\omega_z + \frac{1}{2}(\omega_x^2 + \omega_z^2), \\ \epsilon_{33} &= e_{zz} + e_{zx}\omega_y - e_{yz}\omega_x + \frac{1}{2}(\omega_y^2 + \omega_x^2), \\ \epsilon_{23} &= e_{yz} + \frac{1}{2}\omega_x(e_{zz} - e_{yy}) + \frac{1}{2}\omega_y e_{xy} - \frac{1}{2}\omega_z e_{xx} - \frac{1}{2}\omega_y \omega_z, \\ \epsilon_{31} &= e_{zx} + \frac{1}{2}\omega_y(e_{xx} - e_{zz}) + \frac{1}{2}\omega_z e_{yz} - \frac{1}{2}\omega_x e_{xy} - \frac{1}{2}\omega_z \omega_x, \\ \epsilon_{12} &= e_{xy} + \frac{1}{2}\omega_z(e_{yy} - e_{xx}) + \frac{1}{2}\omega_x e_{zx} - \frac{1}{2}\omega_y e_{yz} - \frac{1}{2}\omega_x \omega_y. \end{aligned} \right\} \quad . \quad . \quad (1.12)$$

At this point it is important to stress the physical significance of these components of strain ϵ . If we look at the homogeneous transformation (1.2) of a small region in the vicinity of a point P attached to the material we now see that it can be obtained as follows :—

(1) We rotate this region as a rigid body. This rotation is defined in first approximation by the vector $\omega_x \omega_y \omega_z$.

(2) A system of rectangular coordinates with point P as origin and parallel with the $x y z$ directions is rigidly rotated by the same amount as the material, and becomes thereby a system which we call (1, 2, 3). With respect to this system 1, 2, 3 we perform the pure deformation (1.6) with the coefficients (1.12).

Therefore, we may also look at the strain components as representing the pure deformation referred to a rectangular frame (1, 2, 3) originally

parallel with the $x y z$ directions and undergoing the same rotation as the material. The strain field is thus referred to a field of rectangular axes whose orientation varies from point to point according to the local rotation of the material. It is important to bear this in mind when we are dealing with the stress, since to correlate stress and strain we must refer them to the same set of axes.

The stress components at a point ξ, η, ζ after deformation referred to the x, y, z directions are denoted by

$$\left. \begin{array}{ccc} \sigma_{xx}, & \sigma_{xy}, & \sigma_{zx}, \\ \sigma_{xy}, & \sigma_{yy}, & \sigma_{yz}, \\ \sigma_{zx}, & \sigma_{yz}, & \sigma_{zz}. \end{array} \right\} \dots \dots \dots (1.13)$$

The stress components relative to the rotated axis (1, 2, 3) are denoted by

$$\left. \begin{array}{ccc} \sigma_{11}, & \sigma_{12}, & \sigma_{31}, \\ \sigma_{12}, & \sigma_{22}, & \sigma_{23}, \\ \sigma_{31}, & \sigma_{23}, & \sigma_{33}. \end{array} \right\} \dots \dots \dots (1.14)$$

In view of further application it is interesting to know the relation between the components (1.13) and (1.14) of the stress.

We have introduced the assumption that the quantities (1.3) are of the first order; we now add the assumption that the *stress components* (1.13) or (1.14) are also of the first order. Since we are interested in a theory of the second order we will drop all terms of higher order than the second in the relation between the stress components (1.13) and (1.14).

In order to do this we may neglect second-order quantities in the expressions for the direction cosines of the axis x, y, z with respect to 1, 2, 3. The transformation formula for the coordinate x, y, z when we rotate the axis into 1, 2, 3 with coordinates x_1, x_2, x_3 are

$$x = x_1 \cos(x_1 1) + x_2 \cos(x_1 2) + x_3 \cos(x_1 3),$$

$$y = x_1 \cos(y_1 1) + x_2 \cos(y_1 2) + x_3 \cos(y_1 3),$$

$$z = x_1 \cos(z_1 1) + x_2 \cos(z_1 2) + x_3 \cos(z_1 3).$$

Now if the rotation of x, y, z into 1, 2, 3 is represented by the vector $\omega_x \omega_y \omega_z$ we have in first approximation

$$x = x_1 - \omega_z x_2 + \omega_y x_3,$$

$$y = \omega_z x_1 + x_2 - \omega_x x_3,$$

$$z = -\omega_y x_1 + \omega_x x_2 + x_3.$$

Therefore the direction cosines are in first approximation

$$\left. \begin{aligned} \cos(x_1 1) &= 1, & \cos(x_1 2) &= -\omega_z, & \cos(x_1 3) &= \omega_y, \\ \cos(y_1 1) &= \omega_z, & \cos(y_1 2) &= 1, & \cos(y_1 3) &= -\omega_x, \\ \cos(z_1 1) &= -\omega_y, & \cos(z_1 2) &= \omega_x, & \cos(z_1 3) &= 1. \end{aligned} \right\} \quad (1.15)$$

We may use these values of the direction cosines in the transformation formulas of the stress components

$$\begin{aligned} \sigma_x &= \sigma_{11} \cos^2(x_1 1) + \sigma_{22} \cos^2(x_1 2) + \sigma_{33} \cos^2(x_1 3) \\ &\quad + 2\sigma_{23} \cos(x_1 2) \cos(x_1 3) + 2\sigma_{31} \cos(x_1 3) \cos(x_1 1) + 2\sigma_{12} \cos(x_1 1) \cos(x_1 2) \\ &\quad \dots \text{etc.} \dots \end{aligned}$$

Neglecting terms of a higher order than the second, we have

$$\left. \begin{aligned} \sigma_{xx} &= \sigma_{11} + 2\sigma_{31}\omega_y - 2\sigma_{12}\omega_z, \\ \sigma_{yy} &= \sigma_{22} + 2\sigma_{12}\omega_z - 2\sigma_{23}\omega_x, \\ \sigma_{zz} &= \sigma_{33} + 2\sigma_{23}\omega_x - 2\sigma_{31}\omega_y, \\ \sigma_{yz} &= \sigma_{23} + (\sigma_{22} - \sigma_{33})\omega_x - \sigma_{12}\omega_y + \sigma_{13}\omega_z, \\ \sigma_{zx} &= \sigma_{31} + (\sigma_{33} - \sigma_{11})\omega_y - \sigma_{23}\omega_z + \sigma_{21}\omega_x, \\ \sigma_{xy} &= \sigma_{12} + (\sigma_{11} - \sigma_{22})\omega_z - \sigma_{31}\omega_x + \sigma_{32}\omega_y. \end{aligned} \right\} \quad (1.16)$$

2. Equilibrium Equations.

A point P of the material originally of coordinates x, y, z acquires the coordinates ξ, η, ζ after deformation. An original closed volume V bounded by the surface S becomes after deformation a volume V' bounded by the surface S'. The x component F_x of the total force acting on the boundary S' after deformation may be expressed by means of a double integral extended to the same material boundary S before deformation. This is done by using the transformation formula of surface integrals. We have

$$\begin{aligned} F_x &= \iint_{S'} \sigma_{xx} d\eta d\zeta + \sigma_{xy} d\zeta d\xi + \sigma_{zx} + \sigma_{xx} d\xi d\eta \\ &= \iint_S \sigma_{xx} \left[\frac{d(\eta, \zeta)}{d(y, z)} dy dz + \frac{d(\eta, \zeta)}{d(z, x)} dz dx + \frac{d(\eta, \zeta)}{d(x, y)} dx dy \right] \\ &\quad + \iint_S \sigma_{xy} \left[\frac{d(\zeta, \xi)}{d(y, z)} dy dz + \frac{d(\zeta, \xi)}{d(z, x)} dz dx + \frac{d(\zeta, \xi)}{d(x, y)} dx dy \right] \\ &\quad + \iint_S \sigma_{zx} \left[\frac{d(\xi, \eta)}{d(y, z)} dy dz + \frac{d(\xi, \eta)}{d(z, x)} dz dx + \frac{d(\xi, \eta)}{d(x, y)} dx dy \right]. \quad (2.1) \end{aligned}$$

In this expression $\frac{d(\eta, \zeta)}{d(y, z)}$, etc. are the partial Jacobians of the transformation of x, y, z into ξ, η, ζ . We have, for instance,

$$\frac{d(\eta, \zeta)}{d(y, z)} = \begin{vmatrix} \frac{\partial \eta}{\partial y} & \frac{\partial \eta}{\partial z} \\ \frac{\partial \zeta}{\partial y} & \frac{\partial \zeta}{\partial z} \end{vmatrix}$$

These Jacobians are the cofactors of the determinant of the differential transformation (1.2).

Since we are interested in a second-order theory, and since these Jacobians appear multiplied by the stress in expression (2.1), we need only keep the linear terms in their values. We find

$$\left. \begin{aligned} \frac{d(\eta, \zeta)}{d(y, z)} &= 1 + e_{yy} + e_{zz} \frac{d(\eta, \zeta)}{d(z, x)} = -(e_{xy} + \omega_z) \frac{d(\eta, \zeta)}{d(x, y)} = -(e_{zx} - \omega_y), \\ \frac{d(\zeta, \xi)}{d(y, z)} &= -(e_{xy} - \omega_z) \frac{d(\zeta, \xi)}{d(z, x)} = 1 + e_{xx} + e_{zz} \frac{d(\zeta, \xi)}{d(x, y)} = -(e_{yz} + \omega_x), \\ \frac{d(\xi, \eta)}{d(y, z)} &= -(e_{zx} + \omega_y) \frac{d(\xi, \eta)}{d(z, x)} = -(e_{yz} - \omega_x) \frac{d(\xi, \eta)}{d(x, y)} = 1 + e_{xx} + e_{yy}. \end{aligned} \right\} \quad (2.2)$$

Introducing these values in (2.1), it becomes

$$F_x = \int_S f_x dS,$$

with

$$\begin{aligned} f_x &= \sigma_{xx}\alpha + \sigma_{xy}\beta + \sigma_{zx}\gamma \\ &\quad [\sigma_{xx}(e_{yy} + e_{zz}) - \sigma_{xy}(e_{xy} - \omega_z) - \sigma_{zx}(e_{zx} + \omega_y)]\alpha \\ &\quad + [-\sigma_{xx}(e_{xy} + \omega_z) + \sigma_{xy}(e_{xx} + e_{zz}) - \sigma_{zx}(e_{yz} - \omega_x)]\beta \\ &\quad + [-\sigma_{xx}(e_{zx} - \omega_y) - \sigma_{xy}(e_{yz} + \omega_x) + \sigma_{zx}(e_{xx} + e_{yy})]\gamma. \end{aligned}$$

This expression f_x is the x component of the force per unit original area, acting at the boundary after deformation, and α, β, γ are the direction cosines of the outside normal to the original boundary S before deformation.

Introducing the component of stress (1.14) with respect to the locally rotated axes 1, 2, 3 through equations (1.16), and dropping terms of order higher than the second,

$$\begin{aligned} f_x &= \sigma_{11}\alpha + \sigma_{12}\beta + \sigma_{31}\gamma \\ &\quad (-\sigma_{12}\omega_z + \sigma_{31}\omega_y)\alpha \\ &\quad (-\sigma_{22}\omega_z + \sigma_{32}\omega_y)\beta \\ &\quad (-\sigma_{23}\omega_z + \sigma_{33}\omega_y)\gamma \end{aligned}$$

$$\begin{aligned}
& [\sigma_{11}(e_{yy} + e_{zz}) - \sigma_{12}e_{xy} - \sigma_{31}e_{xz}]\alpha \\
& [-\sigma_{11}e_{xy} + \sigma_{12}(e_{xx} + e_{zz}) - \sigma_{31}e_{yz}]\beta \\
& [-\sigma_{11}e_{xz} - \sigma_{12}e_{yz} + \sigma_{31}(e_{xx} + e_{yy})]\gamma. \quad . \quad . \quad . \quad . \quad (2.3)
\end{aligned}$$

We have two other similar expressions for f_y and f_z .

Consider now the body force. We assume that the force per unit mass at a point x, y, z is determined by the components $X(x, y, z)$, $Y(x, y, z)$, $Z(x, y, z)$. If ρ is the specific mass before deformation an element of mass $\rho dx dy dz$ keeps the same mass after deformation, but moves to the point ξ, η, ζ . Therefore the x component of the total body force after deformation is

$$\iiint_V X(\xi, \eta, \zeta) \rho dx dy dz. \quad . \quad . \quad . \quad . \quad (2.4)$$

The condition of equilibrium of the volume V after deformation in the x direction is

$$\iint_S f_x dS + \iiint_V X(\xi, \eta, \zeta) \rho dx dy dz = 0.$$

By Green's theorem we change the surface integral into a volume integral extended to V ; we then have

$$\iiint_V [A_x + \rho X(\xi, \eta, \zeta)] dx dy dz = 0.$$

This must be true whatever the volume V , and the equilibrium equations are therefore

$$\begin{aligned}
A_x + \rho X(\xi, \eta, \zeta) = & \frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{12}}{\partial y} + \frac{\partial \sigma_{31}}{\partial z} + \rho X(\xi, \eta, \zeta) \\
& + \frac{\partial}{\partial x} (\sigma_{31}\omega_y) - \frac{\partial}{\partial x} (\sigma_{12}\omega_z) + \frac{\partial}{\partial y} (\sigma_{32}\omega_y) - \frac{\partial}{\partial y} (\sigma_{22}\omega_z) \\
& + \frac{\partial}{\partial z} (\sigma_{33}\omega_y) - \frac{\partial}{\partial z} (\sigma_{23}\omega_z) + \frac{\partial}{\partial x} [\sigma_{11}(e_{yy} + e_{zz})] \\
& + \frac{\partial}{\partial y} [\sigma_{12}(e_{zz} + e_{xx})] + \frac{\partial}{\partial z} [\sigma_{31}(e_{xx} + e_{yy})] \\
& - \frac{\partial}{\partial x} [\sigma_{12}e_{xy} + \sigma_{31}e_{xz}] - \frac{\partial}{\partial y} [\sigma_{11}e_{xy} + \sigma_{31}e_{yz}] \\
& - \frac{\partial}{\partial z} [\sigma_{11}e_{xz} + \sigma_{12}e_{yz}] = 0. \quad . \quad . \quad . \quad . \quad . \quad (2.5)
\end{aligned}$$

There are two other similar equations expressing the equilibrium in the y and z directions. The boundary conditions are determined by expressions (2.3) for the forces f_x, f_y, f_z acting per unit original area of the boundary.

It is possible to give equations (2.5) another remarkable form by taking into account the fact deducible from the equations (2.5) themselves that

$$\begin{aligned}\frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{12}}{\partial y} + \frac{\partial \sigma_{13}}{\partial z} + X\rho, \\ \frac{\partial \sigma_{12}}{\partial x} + \frac{\partial \sigma_{22}}{\partial y} + \frac{\partial \sigma_{23}}{\partial z} + Y\rho, \\ \frac{\partial \sigma_{31}}{\partial x} + \frac{\partial \sigma_{23}}{\partial y} + \frac{\partial \sigma_{33}}{\partial z} + Z\rho\end{aligned}$$

are quantities of the second order and that we have the identities

$$\left. \begin{aligned}\frac{\partial e_{xx}}{\partial z} &= \frac{\partial}{\partial x} [e_{zx} + \omega_y], \\ \frac{\partial e_{yy}}{\partial z} &= \frac{\partial}{\partial y} [e_{yz} - \omega_x], \\ \frac{\partial e_{yy}}{\partial x} &= \frac{\partial}{\partial y} [e_{xy} + \omega_z], \\ \frac{\partial e_{zz}}{\partial x} &= \frac{\partial}{\partial z} [e_{zx} - \omega_y], \\ \frac{\partial e_{zz}}{\partial y} &= \frac{\partial}{\partial z} [e_{yz} + \omega_x], \\ \frac{\partial e_{xx}}{\partial y} &= \frac{\partial}{\partial x} [e_{xy} - \omega_z].\end{aligned}\right\} \dots \dots \dots (2.6)$$

The equilibrium equations (2.5) then take the form

$$\begin{aligned}\frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{12}}{\partial y} + \frac{\partial \sigma_{31}}{\partial z} + \rho X + \rho(\omega_z Y - \omega_y Z) \\ + (\sigma_{11} - \sigma_{22}) \frac{\partial \omega_z}{\partial y} - 2\sigma_{12} \frac{\partial \omega_z}{\partial x} + \sigma_{12} \frac{\partial \omega_x}{\partial z} \\ - (\sigma_{11} - \sigma_{33}) \frac{\partial \omega_y}{\partial z} + 2\sigma_{31} \frac{\partial \omega_y}{\partial x} - \sigma_{31} \\ + \sigma_{32} \frac{\partial \omega_x}{\partial y} \left(\frac{\partial \omega_y}{\partial y} - \frac{\partial \omega_z}{\partial z} \right) \\ + \frac{\partial \sigma_{11}}{\partial x} (e_{yy} + e_{zz}) + \frac{\partial \sigma_{12}}{\partial y} (e_{xx} + e_{zz}) + \frac{\partial \sigma_{13}}{\partial z} (e_{xx} + e_{yy}) \\ - \left(\frac{\partial \sigma_{12}}{\partial x} + \frac{\partial \sigma_{11}}{\partial y} \right) e_{xy} - \left(\frac{\partial \sigma_{31}}{\partial y} + \frac{\partial \sigma_{12}}{\partial z} \right) e_{yz} - \left(\frac{\partial \sigma_{11}}{\partial z} + \frac{\partial \sigma_{31}}{\partial x} \right) e_{zx} = 0, \\ \dots \dots \text{etc.} \dots \dots \dots (2.7)\end{aligned}$$

We may also expand the functions X, Y, Z with respect to u, v, w and write

$$\begin{aligned} X(x+u, y+v, z+w) = & X(x, y, z) + \frac{\partial X}{\partial x} u + \frac{\partial X}{\partial y} v + \frac{\partial X}{\partial z} w, \\ & \dots \text{etc.} \end{aligned} \quad (2.8)$$

Considering X, Y, Z and u, v, w as quantities of the first order, the above expansion includes the first- and second-order terms.

The terms written on the first line of equation (2.7) are the classical terms of the linear theory and additional terms due to the rotation of the material with respect to the body force. The terms on the second and the third line have their origin in relative change of direction and area of opposite faces of an element of material due to its curvature; we call them the *curvature* terms. On the fourth line is what we call the *torsion* term, because it arises from the torsion of an element. On the last two lines are terms depending on the stress gradient and the strain; they will be called the *buoyancy* terms, because they arise from some kind of buoyancy due to the deformation of an element in its own stress field. The curvature and torsion terms vanish in case of a hydrostatic stress condition, while the buoyancy terms vanish when the stress field is homogeneous.

3. The Strain-energy.

The concept of strain-energy and the application of the principle of virtual work throws a new light on the present theory. The first step is to establish a correct expression for the strain-energy. We need only consider a state of homogeneous pure strain defined by the transformation

$$\left. \begin{aligned} \xi &= (1 + \epsilon_{11})x + \epsilon_{12}y + \epsilon_{31}z, \\ \eta &= \epsilon_{12}x + (1 + \epsilon_{22})y + \epsilon_{23}z, \\ \zeta &= \epsilon_{31}x + \epsilon_{23}y + (1 + \epsilon_{33})z. \end{aligned} \right\} \quad (3.1)$$

The homogeneous stress field associated with this deformation is denoted by the constant components

$$\left. \begin{array}{ccc} \sigma_{11}, & \sigma_{12}, & \sigma_{31}, \\ \sigma_{12}, & \sigma_{22}, & \sigma_{23}, \\ \sigma_{31}, & \sigma_{23}, & \sigma_{33}. \end{array} \right\} \quad (3.2)$$

Consider now a cube of unit volume before deformation, its edges being along x, y, z . After deformation it becomes a parallelepiped. On the side originally perpendicular to the z axis acts now a force of component

$$\tau'_{11}, \quad \tau'_{12}, \quad \tau_{13},$$

and on the two other sides will act forces of components

$$\begin{array}{ccc} \tau'_{21}, & \tau'_{22}, & \tau'_{23}, \\ \tau'_{31}, & \tau'_{32}, & \tau'_{33}. \end{array}$$

These forces may be found from the value (2.3) of f_x, f_y, f_z derived above and giving the actual forces at the boundary of a volume. We must introduce ϵ_{11} for e_{xx} , ϵ_{12} for e_{xy} , etc., and put $\omega_x = \omega_y = \omega_z = 0$. We find

$$\begin{aligned} \tau'_{11} &= \sigma_{11}(1 + \epsilon) - \sigma_{11}\epsilon_{11} - \sigma_{12}\epsilon_{12} - \sigma_{31}\epsilon_{31}, \\ \tau'_{12} &= \sigma_{12}(1 + \epsilon) - \sigma_{11}\epsilon_{12} - \sigma_{12}\epsilon_{22} - \sigma_{31}\epsilon_{23}, \\ &\text{etc.}, \end{aligned}$$

$$\begin{aligned} \text{or} \quad \tau'_{\mu\nu} &= \sigma_{\mu\nu}(1 + \epsilon) - \sum_{\alpha}^{\alpha} \sigma_{\mu\alpha}\epsilon_{\nu\alpha}, \quad . \quad . \quad . \quad . \quad . \quad (3.3) \\ \text{with} \quad \epsilon &= \epsilon_{11} + \epsilon_{22} + \epsilon_{33}. \end{aligned}$$

This $\tau'_{\mu\nu}$ is a non-symmetric tensor.

When small increments of strain $\delta\epsilon_{\mu\nu}$ are given to the deformed cube, work is performed by the forces acting on the faces of the parallelepiped. This work, expressing the increment of the strain-energy δW of a unit original volume, is equal to

$$\delta W = \sum^{\mu\nu} \tau'_{\mu\nu} \delta\epsilon_{\mu\nu}. \quad . \quad . \quad . \quad . \quad . \quad (3.4)$$

A more convenient form of δW may be found by writing

$$\delta W = \frac{1}{2} \left[\sum^{\mu\nu} \tau'_{\mu\nu} \delta\epsilon_{\mu\nu} + \sum^{\mu\nu} \tau'_{\nu\mu} \delta\epsilon_{\nu\mu} \right].$$

Since $\delta\epsilon_{\mu\nu} = \delta\epsilon_{\nu\mu}$, we have

$$\delta W = \sum^{\mu\nu} \frac{1}{2} [\tau'_{\mu\nu} + \tau'_{\nu\mu}] \delta\epsilon_{\nu\mu},$$

$$\text{or} \quad \delta W = \sum^{\mu\nu} \tau_{\mu\nu} \delta\epsilon_{\mu\nu}, \quad . \quad . \quad . \quad . \quad . \quad (3.5)$$

$$\begin{aligned} \text{with} \quad \tau_{\mu\nu} &= \frac{1}{2} [\tau'_{\mu\nu} + \tau'_{\nu\mu}], \\ \tau_{\mu\nu} &= (1 + \epsilon) \sigma_{\mu\nu} - \frac{1}{2} \sum_{\alpha}^{\alpha} (\sigma_{\alpha\mu} \epsilon_{\alpha\nu} + \sigma_{\alpha\nu} \epsilon_{\alpha\mu}). \end{aligned} \quad \left. \vphantom{\begin{aligned} \tau_{\mu\nu} &= (1 + \epsilon) \sigma_{\mu\nu} - \frac{1}{2} \sum_{\alpha}^{\alpha} (\sigma_{\alpha\mu} \epsilon_{\alpha\nu} + \sigma_{\alpha\nu} \epsilon_{\alpha\mu}). \end{aligned}} \right\} . \quad . \quad (3.6)$$

The tensor $\tau_{\mu\nu}$ is the symmetric part of $\tau'_{\mu\nu}$. Explicitly, we have

$$\begin{aligned} \tau_{11} &= (1 + \epsilon) \sigma_{11} - \sigma_{11} \epsilon_{11} - \sigma_{12} \epsilon_{12} - \sigma_{31} \epsilon_{31}, \\ \tau_{22} &= (1 + \epsilon) \sigma_{22} - \sigma_{12} \epsilon_{12} - \sigma_{22} \epsilon_{22} - \sigma_{23} \epsilon_{23}, \\ \tau_{33} &= (1 + \epsilon) \sigma_{33} - \sigma_{31} \epsilon_{31} - \sigma_{23} \epsilon_{23} - \sigma_{33} \epsilon_{33}, \\ \tau_{23} &= (1 + \epsilon) \sigma_{23} - \frac{1}{2} (\epsilon_{22} + \epsilon_{33}) \sigma_{23} - \frac{1}{2} (\sigma_{22} + \sigma_{33}) \epsilon_{23} - \frac{1}{2} (\sigma_{12} \epsilon_{31} + \sigma_{31} \epsilon_{12}), \\ \tau_{31} &= (1 + \epsilon) \sigma_{31} - \frac{1}{2} (\epsilon_{33} + \epsilon_{11}) \sigma_{31} - \frac{1}{2} (\sigma_{33} + \sigma_{11}) \epsilon_{31} - \frac{1}{2} (\sigma_{12} \epsilon_{23} + \sigma_{23} \epsilon_{12}), \\ \tau_{12} &= (1 + \epsilon) \sigma_{12} - \frac{1}{2} (\epsilon_{11} + \epsilon_{22}) \sigma_{12} - \frac{1}{2} (\sigma_{11} + \sigma_{22}) \epsilon_{12} - \frac{1}{2} (\sigma_{31} \epsilon_{23} + \sigma_{23} \epsilon_{31}). \end{aligned} \quad (3.6)$$

The symmetric tensor $\tau_{\mu\nu}$ is obviously another representation of the stress. It can easily be seen from the expression of δW that $\tau_{\mu\nu}$ is the symmetric part of the tensor whose components are the *actual forces* acting after deformation on the faces of a parallelepiped which originally was a cube of unit dimension parallel with the coordinate axes. For instance, in the particular case when the coordinate axes are principal directions for both the stress and the strain we have

$$\tau_{11} = (1 + \epsilon_{22} + \epsilon_{33})\sigma_{11},$$

$$\tau_{22} = (1 + \epsilon_{33} + \epsilon_{11})\sigma_{22},$$

$$\tau_{33} = (1 + \epsilon_{11} + \epsilon_{22})\sigma_{33}.$$

In these expressions τ_{11} , for instance, is the stress multiplied by the area $(1 + \epsilon_{22} + \epsilon_{33})$ after deformation of the corresponding face of an original unit cube. The tensor $\tau_{\mu\nu}$ might be also called the *stress per unit initial area before deformation* of the material, while the tensor $\sigma_{\mu\nu}$ would be the *stress per unit actual area after deformation*.

The existence of a potential energy imposes that δW be an exact differential with respect to the strain component. It therefore imposes also certain additional conditions on the stress-strain relations. By these relations $\sigma_{\mu\nu}$ or $\tau_{\mu\nu}$ are expressed as functions of the strain component $\epsilon_{\mu\nu}$. Then the stress-strain relations must satisfy the *fifteen conditions*,

$$\frac{\partial \tau_{\mu\nu}}{\partial \epsilon_{ik}} = \frac{\partial \tau_{ik}}{\partial \epsilon_{\mu\nu}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.7)$$

or

$$\begin{aligned} \frac{\partial \sigma_{\mu\nu}}{\partial \epsilon_{ik}} + \frac{\partial (\epsilon \sigma_{\mu\nu})}{\partial \epsilon_{ik}} - \frac{1}{2} \frac{\partial}{\partial \epsilon_{ik}} \sum^{\alpha} (\sigma_{\alpha\mu} \epsilon_{\alpha\nu} + \sigma_{\alpha\nu} \epsilon_{\alpha\mu}) \\ = \frac{\partial \sigma_{ik}}{\partial \epsilon_{\mu\nu}} + \frac{\partial (\epsilon \sigma_{ik})}{\partial \epsilon_{\mu\nu}} - \frac{1}{2} \frac{\partial}{\partial \epsilon_{\mu\nu}} \sum^{\alpha} (\sigma_{\alpha i} \epsilon_{\alpha k} + \sigma_{\alpha k} \epsilon_{\alpha i}). \quad . \quad . \quad . \quad (3.8) \end{aligned}$$

It is possible to derive the equations of equilibrium (2.7) found above by expressing that the sum of the virtual external and internal work vanishes for all possible variations δu , δv , δw of the coordinates.

The variation of the total strain energy in an original volume V is

$$\delta W = -\delta W_i = \iiint_V^{\mu\nu} \Sigma \tau_{\mu\nu} \delta \epsilon_{\mu\nu} dx dy dz, \quad . \quad . \quad . \quad (3.9)$$

in which we use expressions (1.12) for the strain components $\epsilon_{\mu\nu}$, if $f_x f_y f_z$ denotes the components of the force per unit original area acting at the boundary. The virtual work of the external forces is

$$\begin{aligned} \delta W = \iint_S (f_x \delta u + f_y \delta v + f_z \delta w) dS + \iiint_V [X(\xi, \eta, \zeta) \delta u \\ + Y(\xi, \eta, \zeta) \delta v + Z(\xi, \eta, \zeta) \delta w] \rho dx dy dz. \quad . \quad (3.10) \end{aligned}$$

In the volume integral (3.9) the variations $\delta e_{\mu\nu}$ introduce such quantities as

$$\begin{aligned}\delta e_{xx} &= \delta \frac{\partial u}{\partial x} = \frac{\partial}{\partial x} \delta u, \\ \delta \omega_x &= \frac{1}{2} \delta \left[\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right] = \frac{1}{2} \frac{\partial}{\partial x} \delta v - \frac{1}{2} \frac{\partial}{\partial y} \delta u.\end{aligned}$$

It is therefore possible to bring out the factors δu , δv , δw by partial integration. We find an expression of the form

$$\delta W_i = - \iint_S (p_x \delta u + p_y \delta v + p_z \delta w) dS + \iiint_v (A_x \delta u + A_y \delta v + A_z \delta w) dx dy dz. \quad (3.11)$$

For equilibrium the total virtual work must vanish, hence

$$\begin{aligned}\delta W_i + \delta W_e &= \iiint_v [(A_x + X\rho)\delta u + (A_y + Y\rho)\delta v + (A_z + Z\rho)\delta w] dx dy dz \\ &\quad + \iint_S [(f_x - p_x)\delta u + (f_y - p_y)\delta v + (f_z - p_z)\delta w] dS = 0. \quad (3.12)\end{aligned}$$

This must be identically zero for arbitrary values of δu , δv , δw , therefore we have the conditions

$$\begin{aligned}A_x + X\rho &= \frac{\partial \tau_{11}}{\partial x} + \frac{\partial \tau_{12}}{\partial y} + \frac{\partial \tau_{13}}{\partial z} + \rho X(\xi, \eta, \zeta) + \frac{\partial}{\partial x} (\tau_{31} \omega_y) - \frac{\partial}{\partial x} (\tau_{12} \omega_z) \\ &\quad + \frac{\partial}{\partial y} (\tau_{23} \omega_y) - \frac{\partial}{\partial y} (\tau_{22} \omega_z) + \frac{\partial}{\partial z} (\tau_{33} \omega_y) - \frac{\partial}{\partial z} (\tau_{23} \omega_z) \\ &\quad + \frac{1}{2} \frac{\partial}{\partial y} [(\tau_{22} - \tau_{11})e_{xy}] + \frac{1}{2} \frac{\partial}{\partial z} [(\tau_{33} - \tau_{11})e_{xz}] \\ &\quad - \frac{1}{2} \frac{\partial}{\partial y} [\tau_{12}(e_{yy} - e_{xx})] - \frac{1}{2} \frac{\partial}{\partial z} [\tau_{31}(e_{zz} - e_{xx})] \\ &\quad + \frac{1}{2} \frac{\partial}{\partial y} [\tau_{23}e_{zx} - \tau_{31}e_{yz}] + \frac{1}{2} \frac{\partial}{\partial z} [\tau_{23}e_{xy} - \tau_{12}e_{yz}] = 0, \quad (3.13)\end{aligned}$$

and two other equations as above :

$$\begin{aligned}A_y + Y\rho &= 0, \\ A_z + Z\rho &= 0.\end{aligned}$$

These are equilibrium equations expressed by means of the stress $\tau_{\mu\nu}$ referred to the original areas before deformation.

We also derive the boundary conditions

$$\begin{aligned}f_x = p_x &= \tau_{11}\alpha + \tau_{12}\beta + \tau_{31}\gamma + (\tau_{31}\omega_y - \tau_{12}\omega_z)\alpha + (\tau_{23}\omega_y - \tau_{22}\omega_z)\beta \\ &\quad + (\tau_{33}\omega_y - \tau_{23}\omega_z)\gamma + [\tfrac{1}{2}(\tau_{22} - \tau_{11})e_{xy} - \tfrac{1}{2}\tau_{12}(e_{yy} - e_{xx})] \\ &\quad + \tfrac{1}{2}(\tau_{23}e_{zx} - \tau_{31}e_{yz})\beta + [\tfrac{1}{2}(\tau_{33} - \tau_{11})e_{xz} - \tfrac{1}{2}\tau_{31}(e_{zz} - e_{xx}) \\ &\quad + \tfrac{1}{2}(\tau_{23}e_{xy} - \tau_{12}e_{yz})]\gamma, \quad (3.14)\end{aligned}$$

and two other similar equations,

$$f_y = p_y,$$

$$f_z = p_z.$$

We may now introduce the stresses $\sigma_{\mu\nu}$ per unit area after deformation by substituting in the above equation the values (3.6) for $\tau_{\mu\nu}$ and dropping all terms of higher order than the second. Equations (3.13) then become

$$\begin{aligned} & \frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{12}}{\partial y} + \frac{\partial \sigma_{31}}{\partial z} + \rho X(\xi, \eta, \zeta) + \frac{\partial}{\partial x}(\sigma_{31}\omega_y) - \frac{\partial}{\partial y}(\sigma_{12}\omega_z) \\ & + \frac{\partial}{\partial y}(\sigma_{32}\omega_y) - \frac{\partial}{\partial y}(\sigma_{22}\omega_z) + \frac{\partial}{\partial z}(\sigma_{33}\omega_y) - \frac{\partial}{\partial z}(\sigma_{23}\omega_z) \\ & \times \frac{\partial}{\partial x}[\sigma_{11}(e_{yy} + e_{zz})] + \frac{\partial}{\partial y}[\sigma_{12}(e_{zz} + e_{xx})] + \frac{\partial}{\partial z}[(\sigma_{31}(e_{xx} + e_{yy}))] \\ & - \frac{\partial}{\partial x}[\sigma_{12}e_{xy} + \sigma_{31}e_{zx}] - \frac{\partial}{\partial y}[\sigma_{11}e_{xy} + \sigma_{31}e_{yz}] \\ & - \frac{\partial}{\partial z}[\sigma_{11}e_{zx} + \sigma_{12}e_{yz}] = 0, \\ & \text{etc.,} \end{aligned} \quad (3.15)$$

and the boundary conditions (3.14) become

$$\begin{aligned} f_x = & \sigma_{11}\alpha + \sigma_{12}\beta + \sigma_{31}\gamma + (\sigma_{31}\omega_y - \sigma_{12}\omega_z)\alpha + (\sigma_{23}\omega_y - \sigma_{22}\omega_z)\beta \\ & + (\sigma_{33}\omega_y - \sigma_{23}\omega_z)\gamma + \sigma_{11}(e_{yy} + e_{zz})\alpha + \sigma_{12}(e_{xx} + e_{zz})\beta + \sigma_{31}(e_{xx} + e_{zz})\gamma \\ & - (\sigma_{12}e_{xy} + \sigma_{31}e_{zx})\alpha - (\sigma_{11}e_{xy} + \sigma_{31}e_{yz})\beta - (\sigma_{11}e_{zx} + \sigma_{12}e_{yz})\gamma, \\ & \text{etc.} \end{aligned} \quad (3.16)$$

These equilibrium equations and boundary conditions are identical with those (2.7) and (2.3) found above by a different method.

4. Material under Initial Stress.

The previous methods may be readily applied to establish a linear theory of elasticity for small deformations in a material under initial stress. The initial state with coordinates $x y z$ is here associated with initial stresses

$$\left. \begin{array}{ccc} S_{xx}, & S_{xy}, & S_{zx}, \\ S_{xy}, & S_{yy}, & S_{yz}, \\ S_{zx}, & S_{yz}, & S_{zz}. \end{array} \right\} \quad (4.1)$$

These initial stresses, being in equilibrium, satisfy the following equations :

$$\left. \begin{aligned} \frac{\partial S_{xx}}{\partial x} + \frac{\partial S_{xy}}{\partial y} + \frac{\partial S_{xz}}{\partial z} + \rho X &= 0, \\ \frac{\partial S_{xy}}{\partial x} + \frac{\partial S_{yy}}{\partial y} + \frac{\partial S_{yz}}{\partial z} + \rho Y &= 0, \\ \frac{\partial S_{xz}}{\partial x} + \frac{\partial S_{yz}}{\partial y} + \frac{\partial S_{zz}}{\partial z} + \rho Z &= 0. \end{aligned} \right\} \dots \dots \dots (4.2)$$

If now the material undergoes small deformations, so that the initial coordinates $x y z$ become $\xi = x + u$, $\eta = y + v$, $\zeta = z + w$, the stresses undergo slight changes. Referring the stress to axes 1, 2, 3 rotating locally with the material through an amount defined by ω_x , ω_y , ω_z the stresses become

$$\sigma_{\mu\nu} = \left\{ \begin{array}{ccc} S_{xx} + s_{11}, & S_{xy} + s_{12}, & S_{xz} + s_{31}, \\ S_{xy} + s_{12}, & S_{yy} + s_{22}, & S_{yz} + s_{23}, \\ S_{xz} + s_{31}, & S_{yz} + s_{23}, & S_{zz} + s_{33}. \end{array} \right\} \dots \dots (4.3)$$

Now we are interested here in a linear theory with respect to u , v , w , the strain, the rotation, and the stress increments, which quantities are all assumed to be small of the first order.

We adopt, therefore, for the components of the strain ϵ the first-order approximation

$$\epsilon = \left\{ \begin{array}{ccc} e_{xx}, & e_{xy}, & e_{zx}, \\ e_{xy}, & e_{yy}, & e_{yz}, \\ e_{zx}, & e_{yz}, & e_{zz}. \end{array} \right\} \dots \dots \dots (4.4)$$

Because the local reference axes 1, 2, 3 rotate with the material the stress increments

$$s = \left\{ \begin{array}{ccc} s_{11}, & s_{12}, & s_{31}, \\ s_{12}, & s_{22}, & s_{23}, \\ s_{31}, & s_{23}, & s_{33}. \end{array} \right\} \dots \dots \dots (4.5)$$

depend only on the strain ϵ . These stress-strain relations may be taken linear in first-order approximation. However, some remark will have to be made later regarding the coefficients, as the properties of the latter are not the same as in the case of no initial stress.

The additional stress may be due to a change in boundary forces Δf_x , Δf_y , Δf_z or in the volume forces ΔX , ΔY , ΔZ , or both. These increments are also considered as first-order quantities. We may proceed exactly along the same lines as in the previous non-linear theory, in which the stress components (1.14) will be replaced by expression (4.3), and then drop in the formulas all the terms which are not linear with

respect to the first-order quantities. For instance, the stress component referred to initial directions $x y z$ is given by expressions (1.16), in which we substitute the stress components (4.3) and drop all second-order terms. We find

$$\left. \begin{aligned} \sigma_{xx} &= S_{xx} + s_{11} + 2S_{zx}\omega_y - 2S_{xy}\omega_z, \\ \sigma_{yy} &= S_{yy} + s_{22} + 2S_{xy}\omega_z - 2S_{yz}\omega_x, \\ \sigma_{zz} &= S_{zz} + s_{33} + 2S_{yz}\omega_x - 2S_{zx}\omega_y, \\ \sigma_{yz} &= S_{yz} + s_{23} + (S_{yy} - S_{zz})\omega_x - S_{xy}\omega_y + S_{xz}\omega_z, \\ \sigma_{zx} &= S_{zx} + s_{31} + (S_{zz} - S_{xx})\omega_y - S_{yz}\omega_z + S_{yx}\omega_x, \\ \sigma_{xy} &= S_{xy} + s_{12} + (S_{xx} - S_{yy})\omega_z - S_{zx}\omega_x + S_{zy}\omega_y. \end{aligned} \right\} \quad (4.6)$$

Similarly the same substitution of the components (4.3) instead of σ (1.14) in equations (2.7), and taking into account the initial equilibrium conditions (4.2), yields the following equilibrium conditions

$$\begin{aligned} \frac{\partial s_{11}}{\partial x} + \frac{\partial s_{12}}{\partial y} + \frac{\partial s_{13}}{\partial z} + \rho \left(\frac{\partial X}{\partial x} u + \frac{\partial X}{\partial y} v + \frac{\partial X}{\partial z} w \right) + \rho(\omega_z Y - \omega_y Z) \\ + (S_{zz} - S_{yy}) \frac{\partial \omega_z}{\partial y} - 2S_{xy} \frac{\partial \omega_z}{\partial x} + S_{xy} \frac{\partial \omega_x}{\partial z} - (S_{xx} - S_{zz}) \frac{\partial \omega_y}{\partial z} \\ - 2S_{zx} \frac{\partial \omega_y}{\partial x} - S_{zx} \frac{\partial \omega_x}{\partial y} + S_{zy} \left(\frac{\partial \omega_y}{\partial y} - \frac{\partial \omega_z}{\partial z} \right) + \frac{\partial S_{xx}}{\partial z} (e_{yy} + e_{zz}) \\ + \frac{\partial S_{xy}}{\partial y} (e_{xx} + e_{zz}) + \frac{\partial S_{xz}}{\partial z} (e_{xx} + e_{yy}) - \left(\frac{\partial S_{xy}}{\partial x} + \frac{\partial S_{xx}}{\partial y} \right) e_{xy} \\ - \left(\frac{\partial S_{xz}}{\partial y} + \frac{\partial S_{xy}}{\partial z} \right) e_{yz} - \left(\frac{\partial S_{xx}}{\partial z} + \frac{\partial S_{xy}}{\partial x} \right) e_{zx} = 0, \\ \text{etc.} \end{aligned} \quad (4.7)$$

The boundary conditions for the increment of boundary force Δf are

$$\begin{aligned} \Delta f_x &= s_{11}\alpha + s_{12}\beta + s_{31}\gamma + [S_{zx}\omega_y - S_{xy}\omega_z]\alpha + [S_{yz}\omega_y - S_{yy}\omega_z]\beta \\ &+ S_{zz}\omega_y - S_{yz}\omega_z\gamma + S_{xx}[e_{yy} + e_{zz}]\alpha + S_{xy}(e_{xx} + e_{zz})\beta + S_{zx}(e_{xx} + e_{yy})\gamma \\ &- (S_{xy}e_{xy} + S_{zx}e_{zx})\alpha - (S_{xx}e_{xy} + S_{zx}e_{yz})\beta - (S_{xx}e_{zx} + S_{xy}e_{yz})\gamma. \end{aligned} \quad (4.8)$$

We recognize here in equation (4.7) terms of the same physical nature as in equation (2.7) of the non-linear theory. The curvature has no effect when the initial stress is hydrostatic. When the initial stress field is homogeneous the buoyancy terms disappear, and we are left with those of the first line and the curvature terms. It can be seen that the curvature terms are those playing the fundamental rôle in buckling phenomena.

We may also refer the stresses to the original areas, *i. e.*, use the stress

tensor $\tau_{\mu\nu}$ instead of $\sigma_{\mu\nu}$. The stress components before the deformation are the same as above (4.1). After deformation the stress $\tau_{\mu\nu}$ is

$$\tau_{\mu\nu} = \begin{Bmatrix} S_{xx} + t_{11}, & S_{xy} + t_{12}, & S_{zx} + t_{31}, \\ S_{xy} + t_{12}, & S_{yy} + t_{22}, & S_{yz} + t_{23}, \\ S_{zx} + t_{31}, & S_{yz} + t_{23}, & S_{zz} + t_{33}. \end{Bmatrix} \quad (4.9)$$

However, the stress increments $t_{\mu\nu}$ are not the same as the increments $s_{\mu\nu}$ used before. In fact from relation (3.6) above we derive

$$\tau_{\mu\nu} = s_{\mu\nu} + S_{\mu\nu}e - \frac{1}{2} \sum^{\alpha} (S_{\alpha\mu}e_{\alpha\nu} + S_{\alpha\nu}e_{\alpha\mu}), \\ e = e_{xx} + e_{yy} + e_{zz}.$$

Explicitly

$$\left. \begin{aligned} t_{11} &= s_{11} + eS_{xx} - S_{xx}e_{xx} - S_{xy}e_{xy} - S_{zx}e_{zx}, \\ t_{22} &= s_{22} + eS_{yy} - S_{xy}e_{xy} - S_{yz}e_{yz} - S_{yz}e_{yz}, \\ t_{33} &= s_{33} + eS_{zz} - S_{zx}e_{zx} - S_{yz}e_{yz} - S_{zz}e_{zz}, \\ t_{23} &= s_{23} + eS_{yz} - \frac{1}{2}(e_{yy} + e_{zz})S_{yz} - \frac{1}{2}(S_{yy} + S_{zz})e_{yz} - \frac{1}{2}(S_{yz}e_{zx} + S_{zx}e_{xy}), \\ t_{31} &= s_{31} + eS_{zx} - \frac{1}{2}(e_{zz} + e_{xx})S_{zx} - \frac{1}{2}(S_{zz} + S_{xx})e_{zx} - \frac{1}{2}(S_{xy}e_{yz} + S_{yz}e_{xy}), \\ t_{12} &= s_{12} + eS_{xy} - \frac{1}{2}(e_{xx} + e_{yy})S_{xy} - \frac{1}{2}(S_{xx} + S_{yy})e_{xy} - \frac{1}{2}(S_{zx}e_{yz} + S_{yz}e_{zx}). \end{aligned} \right\} \quad (4.10)$$

The equilibrium equations with these components are

$$\begin{aligned} \frac{\partial t_{11}}{\partial x} + \frac{\partial t_{12}}{\partial y} + \frac{\partial t_{13}}{\partial z} + \rho \Delta x + \rho \left(\frac{\partial X}{\partial x} u + \frac{\partial X}{\partial y} v + \frac{\partial X}{\partial z} w \right) + \frac{\partial}{\partial x} (S_{xx}\omega_y) \\ - \frac{\partial}{\partial x} (S_{xy}\omega_z) + \frac{\partial}{\partial y} (S_{yz}\omega_y) - \frac{\partial}{\partial y} (S_{yy}\omega_z) + \frac{\partial}{\partial z} (S_{zz}\omega_y) - \frac{\partial}{\partial z} (S_{yz}\omega_z) \\ + \frac{1}{2} \frac{\partial}{\partial y} [S_{yy} - S_{xx})e_{xy}] + \frac{1}{2} \frac{\partial}{\partial z} [(S_{zz} - S_{xx})e_{zx}] \\ - \frac{1}{2} \frac{\partial}{\partial y} [S_{xy}(e_{yy} - e_{xx})] - \frac{1}{2} \frac{\partial}{\partial z} [S_{zx}(e_{zz} - e_{xx})] \\ + \frac{1}{2} \frac{\partial}{\partial y} [S_{yz}e_{zx} - S_{zx}e_{yz}] + \frac{1}{2} \frac{\partial}{\partial z} [S_{yz}e_{xy} - S_{xy}e_{yz}] = 0. \quad (4.11) \end{aligned}$$

We also have the boundary condition for the increment f of the boundary force

$$\begin{aligned} \Delta f_x &= t_{11}\alpha + t_{12}\beta + t_{31}\gamma + (S_{zz}\omega_y - S_{xy}\omega_z)\alpha + S_{yz}\omega_y - S_{yy}\omega_z\beta \\ &\quad + (S_{zz}\omega_y - S_{yz}\omega_z)\gamma + [\frac{1}{2}(S_{yy} - S_{xx})e_{xy} - \frac{1}{2}(e_{yy} - e_{xx})S_{xy} \\ &\quad + \frac{1}{2}(S_{yz}e_{zx} - S_{zx}e_{yz})]\beta + [\frac{1}{2}(S_{zz} - S_{xx})e_{zx} - \frac{1}{2}(e_{zz} - e_{xx})S_{zx} \\ &\quad + \frac{1}{2}(S_{yz}e_{xy} - S_{xy}e_{yz})]\gamma, \\ &\text{etc.} \quad (4.12) \end{aligned}$$

These equations are readily verified to be equivalent to those found by C. B. Biezeno and H. Hencky⁽⁴⁾, but derived by an entirely different method.

The same result may be derived from the strain-energy viewpoint. The strain-energy corresponding to a linear theory must contain both the linear and the quadratic terms; therefore we must use here the second-order approximation (1.12) for the strain $\epsilon_{\mu\nu}$. We substitute the component (4.9) for $\tau_{\mu\nu}$ in the potential energy variation (3.5). We find

$$\delta W = \sum^{\mu\nu} t_{\mu\nu} \delta \epsilon_{\mu\nu} + \sum^{\mu\nu} S_{\mu\nu} \delta \epsilon_{\mu\nu}. \quad (4.13)$$

The notation $S_{11}, S_{12} \dots$ etc. is used for $S_{xx}, S_{xy} \dots$ etc.

We have assumed that the stress increments are linear functions of the strains

$$t_{\mu\nu} = \sum^{kr} C_{\mu\nu}^{kr} \epsilon_{rk}. \quad (4.14)$$

The sum is extended to all six combinations of k and r . The assumption that there exists a strain-energy implies that δW is an exact differential in $\delta \epsilon_{\mu\nu}$, hence that

$$\frac{\partial t_{\mu\nu}}{\partial \epsilon_{kr}} = \frac{\partial t_{rk}}{\partial \epsilon_{\mu\nu}} \quad \text{or} \quad C_{\mu\nu}^{kr} = C_{kr}^{\mu\nu}. \quad (4.15)$$

This implies fifteen relations between the coefficients of the stress-strain relation expressed by the fact that the matrix of the coefficients of the stress-strain relation is symmetric. However, it is important to notice that this only holds when we use the $t_{\mu\nu}$ stress components, i. e., the stresses per unit original area before deformation.

If we use the actual stresses $s_{\mu\nu}$ referred to the area after deformation we may write the linear stress-strain relations in the form

$$s_{\mu\nu} = \sum^{kr} B_{\mu\nu}^{kr} \epsilon_{rk}. \quad (4.16)$$

The stress s is related to the stress t by relations (4.10), in which we may indifferently write ϵ instead of e . Using these relations the conditions (4.15) above become

$$\begin{aligned} B_{\mu\nu}^{kr} + S_{\mu\nu} \frac{\partial \epsilon}{\partial \epsilon_{\mu\nu}} - \frac{1}{2} \frac{\partial}{\partial \epsilon_{kr}} \sum^{\alpha} (S_{\alpha\mu} \epsilon_{\alpha\nu} + S_{\alpha\nu} \epsilon_{\alpha\mu}) \\ = B_{kr}^{\mu\nu} + S_{kr} \frac{\partial \epsilon}{\partial \epsilon_{\mu\nu}} - \frac{1}{2} \frac{\partial}{\partial \epsilon_{\mu\nu}} \sum^{\alpha} (S_{\alpha k} \epsilon_{\alpha r} + S_{\alpha r} \epsilon_{\alpha k}). \end{aligned} \quad (4.17)$$

We can see that in general the coefficients B do not constitute a symmetric matrix. For instance, in two dimensions the stress-strain relations (4.14) for the stress $t_{\mu\nu}$ referred to the original area is

$$\begin{aligned} t_{11} &= C_{11}^{11} \epsilon_{11} + C_{11}^{22} \epsilon_{22} + C_{11}^{12} \epsilon_{12}, \\ t_{22} &= C_{22}^{11} \epsilon_{11} + C_{22}^{22} \epsilon_{22} + C_{22}^{12} \epsilon_{12}, \\ t_{12} &= C_{12}^{11} \epsilon_{11} + C_{12}^{22} \epsilon_{22} + C_{12}^{12} \epsilon_{12}, \end{aligned}$$

with

$$\begin{aligned}C_{22}^{11} &= C_{11}^{22}, \\ C_{12}^{11} &= C_{11}^{12}, \\ C_{12}^{22} &= C_{22}^{12}.\end{aligned}$$

The stress $s_{\mu\nu}$ referred to the actual areas after deformation is related to $t_{\mu\nu}$ by relation (4.10). They are in two dimensions

$$\begin{aligned}t_{11} &= s_{11} + S_{11}\epsilon_{22} - S_{12}\epsilon_{12}, \\ t_{22} &= s_{22} + S_{22}\epsilon_{11} - S_{12}\epsilon_{12}, \\ t_{12} &= s_{12} + \tfrac{1}{2}(\epsilon_{11} + \epsilon_{22})S_{12} - \tfrac{1}{12}(S_{11} + S_{22})\epsilon_{12}.\end{aligned}$$

The stress-strain relations for the stress s are

$$\begin{aligned}s_{11} &= B_{11}^{11}\epsilon_{11} + B_{11}^{22}\epsilon_{22} + B_{11}^{12}\epsilon_{12}, \\ s_{22} &= B_{22}^{11}\epsilon_{11} + B_{22}^{22}\epsilon_{22} + B_{22}^{12}\epsilon_{12}, \\ s_{12} &= B_{12}^{11}\epsilon_{11} + B_{12}^{22}\epsilon_{22} + B_{12}^{12}\epsilon_{12}.\end{aligned}$$

The coefficients B must satisfy the relations (4.17), which become in this case

$$\begin{aligned}B_{11}^{12} + S_{11} &= B_{12}^{11} + S_{22}, \\ B_{11}^{12} - S_{12} &= B_{12}^{11} + \tfrac{1}{2}S_{12}, \\ B_{22}^{12} - S_{12} &= B_{12}^{22} + \tfrac{1}{2}S_{12}.\end{aligned}$$

The coefficients in this case will in general not be symmetric except when the initial stress is a hydrostatic pressure

$$S_{11} = S_{22}, \quad S_{12} = 0.$$

Let us now go back to the expression (4.13) for the variation of potential energy δW . The quantity $\delta W' = \sum t_{\mu\nu} \delta \epsilon_{\mu\nu}$ is the differential of an homogeneous quadratic form W' in $\epsilon_{\mu\nu}$. We have

$$\frac{\partial W'}{\partial \epsilon_{\mu\nu}} = t_{\mu\nu}.$$

According to Euler's theorem for homogeneous forms

$$W' = \tfrac{1}{2} \sum \frac{\partial W'}{\partial \epsilon_{\mu\nu}} \epsilon_{\mu\nu} = \tfrac{1}{2} \sum^{\mu\nu} t_{\mu\nu} \epsilon_{\mu\nu}$$

Hence the expressions for the strain-energy in a body under initial stress

$$W = \tfrac{1}{2} \sum^{\mu\nu} t_{\mu\nu} \epsilon_{\mu\nu} + \sum^{\mu\nu} S_{\mu\nu} \epsilon_{\mu\nu}. \quad . \quad . \quad . \quad . \quad . \quad (4.18)$$

The value of the strain $\epsilon_{\mu\nu}$ must be that given by formulas (1.12) above. However, in the quadratic expressions $\sum t_{\mu\nu} \epsilon_{\mu\nu}$ we may substitute $e_{\mu\nu}$ for $\epsilon_{\mu\nu}$, as this does not affect the second-order terms, and write

$$W = \tfrac{1}{2} \sum^{\mu\nu} t_{\mu\nu} e_{\mu\nu} + \sum^{\mu\nu} S_{\mu\nu} \epsilon_{\mu\nu}. \quad . \quad . \quad . \quad . \quad . \quad (4.19)$$

with the boundary conditions

$$f(x) = \tau_{11}\alpha + \tau_{12}\beta + \tau_{31}\gamma + (\tau_{31}\omega_y - \tau_{12}\omega_z)\alpha \\ + (\tau_{23}\omega_y - \tau_{22}\omega_z) + \beta(\tau_{33}\omega_y - \tau_{23}\omega_z)\gamma.$$

Similar equations are found for the linearized theory of a body under high initial stress. Such equations for the two-dimensional case have been derived already in 1934 by the author⁽⁵⁾.

We have written the above equations with the stress components $\tau_{\mu\nu}$ referred to the initial areas. We may substitute in these equations the values (3.6) of $\tau_{\mu\nu}$ in terms of $\sigma_{\mu\nu}$.

References.

- (1) See, for instance, Murnaghan, "Finite Deformations of an Elastic Solid," American Journal of Mathematics, April 1937.
- (2) R. V. Southwell, "On the General Theory of Elastic Stability," Phil. Trans. Roy. Soc. ser. A, cexiii. pp. 187-244 (1913).
- (3) E. Trefftz, "Zur Theorie der Stabilität des Elastischen Gleichgewichts," Zeit. für Ang. Math. u. Mech. Bd. xii. Heft 2, pp. 160-165 (April 1933).
- (4) C. B. Biezeno and H. Hencky, 'Proceedings of the Royal Academy, Amsterdam,' xxxi. n. 6.
- (5) M. A. Biot, "Sur la Stabilité de l'Equilibre Elastique—Equations de l'Elasticité d'un milieu soumis à tension initiale," Ann. Soc. Scient. de Bruxelles, tome liv. ser. B, p. 18 (1934).
- (6) A. E. H. Love, 'Mathematical Theory of Elasticity,' p. 65. Cambridge (1906).

XLIV. *Vector Maps of Finite and Periodic Point Sets.*

By D. M. WRINCH, M.A., D.Sc., Mathematical Institute, Oxford *.

[Received February 21, 1939.]

1. *Introduction.*

A SET of points P_n in space S_1 is given, with each of which a corresponding intensity r_n is associated. To construct the vector map in space S_2 of the point intensity set $\Sigma r_n P_n$, we select any point O_2 as origin in S_2 and erect at O_2 vectors corresponding to vector distances from P_n to P_m , and from P_m to P_n , for all values of n and m including $n=m$, associating with $P_n P_m$ and $P_m P_n$, the end points of these vectors, intensities $r_n r_m$. Corresponding to any set of point intensities in S_1 there is thus a second set of point intensities in S_2 which constitute its vector map. The question then arises as to the relation between point intensity sets in S_1 and in vector space S_2 . Given the S_1 set, what is the S_2 set? Further, what is the complete inventory of S_1 sets with a given S_2 set as vector map?

In a previous communication † certain aspects of the geometry of the vector maps of finite sets have been considered. On the present occasion such maps are investigated analytically, with particular reference to the structure of the hexadic vector maps belonging to triadic sets. Special attention is also given to the transition between vector maps of finite and periodic point sets and to the correlation between periodic point sets and their periodic vector maps.

2. *Vector Maps of Finite Point Sets.*

We consider first a simple point intensity distribution in S_1 .

Suppose in space S_1 there are six coplanar points $P_1, P_2, P_3, P_4, P_5, P_6$ with intensities $r_1, r_2, r_3, r_4, r_5, r_6$ respectively, then the distribution in S_1 is

$$A = r_1 P_1 + r_2 P_2 + r_3 P_3 + r_4 P_4 + r_5 P_5 + r_6 P_6.$$

We then define the distribution in S_2 by the vector function

$$V(A) = \sum_n r_n^2 \cdot P_n P_n + \sum_{m \neq n} r_n r_m P_n P_m,$$

where $P_r P_s$ is the end point of the vector from P_r to P_s erected at O_2 , the origin in S_2 , and in particular $P_r P_r$ lies at O_2 itself. This distribution

* Communicated by the Author.

† Phil. Mag. xxvii. p. 98 (1939).

necessarily has a centre of symmetry and comprises, in addition to a point of intensity Σr_n^2 at the origin O_2 , 15 pairs of points. Of these 15 pairs, six pairs are due to interactions of adjacent points in S_1 , such as (P_1, P_2) , (P_2, P_3) , . . . , six pairs to interactions of points next but one, such as (P_1, P_3) , (P_2, P_4) , . . . , and three pairs to interactions of points next but two, namely (P_1, P_4) , (P_2, P_5) and (P_3, P_6) . If the hexagon $P_1P_2P_3P_4P_5P_6$ has no symmetry, then the S_2 distribution has no symmetries other than a centre of symmetry.

Let us now proceed to treat the matter formally. We confine ourselves to two-dimensional distributions in S_1 and consequently consider only two-dimensional distributions in S_2 . We express points by means of their polar coordinates. Thus a point of intensity r , whose distance from the origin O_1 is R and whose vectorial angle is ϕ namely the point $R \exp i\phi$ will be written $r(R, \phi)$. The general distribution in atomic space is then

$$A = c(0, -) + \Sigma r_n(R_n, \phi_n).$$

In the first instance we treat only finite distributions, so that n runs through a finite number of integer values, 1, 2, 3, . . . N . When convenient we talk of "the point n " and "the point intensity n ," meaning by these terms the point (R_n, ϕ_n) and the point intensity $r_n(R_n, \phi_n)$ respectively.

Now the vector from the point m to the point n in S_1 gives in S_2 the point (R_{mn}, ϕ_{mn}) , where

$$R_{mn} \exp i\phi_{mn} = R_n \exp i\phi_n - R_m \exp i\phi_m,$$

so that

$$R_{mn}^2 = R_n^2 + R_m^2 - 2R_n R_m \cos(\phi_n - \phi_m)$$

and $\tan \phi_{nm} = (R_n \sin \phi_n - R_m \sin \phi_m) / (R_n \cos \phi_n - R_m \cos \phi_m)$.

Thus the interaction of two point intensities n and m in S_1 , gives in S_2 two point intensities, namely, $r_n r_m (R_{nm}, \phi_{nm})$ and $r_n r_m (R_{nm}, \phi_{nm} + \pi)$, or, as it will be more convenient to write,

$$r_n r_m (R_{nm}, \phi_{nm})_2,$$

meaning by this symbol, the point intensity "dyad" centred at O_2 . Since the interaction of any two points gives a dyad centred at O_2 , every vector map has a centre of symmetry. Since (R_n, ϕ_n) is the point O_2 , therefore every vector map has at O_2 a point intensity which may be written $(c^2 + \Sigma r_n^2)(0, -)$. Thus corresponding to the point intensity distribution A above, the vector function is

$$V(A) = (c^2 + \Sigma r_n^2)(0, -) + \Sigma r_n r_m (R_{nm}, \phi_{nm})_2.$$

Evidently every vector function V is invariant for a change in sign throughout of every r_n , and when every point in A undergoes a parallel displacement: also for a rotation of π about any coplanar point. Further,

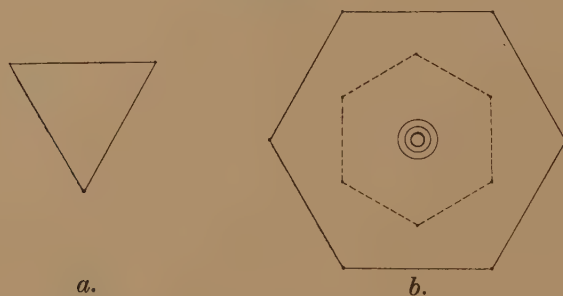
the effect of rotating every term in A through ψ is to rotate $V(A)$ also through ψ . V is unchanged when $\psi = \pi$, since V has a centre of symmetry.

In the general case, when the A distribution has no symmetries, this expression cannot be simplified.

3. Trigonal Distributions in S_1 .

If ABC are the vertices of an equilateral triangle whose centre is O_1 , the origin of coordinates, then A, B, C form a point triad, denoted by $(R, \phi)_3$, where (R, ϕ) are the polar coordinates of A, B , or C . If to each of these points the same intensity r is assigned, the three point intensities then form a point intensity triad $r(R, \phi)_3$. Similarly, the vertices of a central regular hexagon, one of whose points has polar coordinates (R, ϕ) , form a point hexad denoted by $(R, \phi)_6$, and a point intensity hexad $r(R, \phi)_6$ results if to each of these six points the same intensity r is assigned.

Fig. 1.



The most general intensity distribution which is trigonal, is made up of any number of the concentric intensity triads, together with an intensity c at the common centre O_1 , namely,

$$A = c(0, -) + r_1(R_1, \phi_1)_3 + r_2(R_2, \phi_2)_3 + \dots + r_n(R_n, \phi_n)_3.$$

The domestic interactions of the intensity triad $r(R, \phi)_3$ gives (fig. 1)

$$3r^2(0, -) + r^2(\sqrt{3}R, \phi + \pi/6)_6.$$

The addition of the point intensity c at the origin gives, in addition (fig. 1), $cr(R, \phi)_6 + c^2(0, -)$. For the interaction of two different triads $r_m(R_m, \phi_m)_3$ and $r_n(R_n, \phi_n)_3$, we find three hexads,

$$r_m r_n (R_{mn}, \phi_{mn})_6 + r_m r_n (R_{mn}', \phi_{mn}')_6 + r_m r_n (R_{mn}'', \phi_{mn}'')_6,$$

where

$$(R_{mn}, \phi_{mn}) = (R_n, \phi_n) - (R_m, \phi_m),$$

$$(R_{mn}', \phi_{mn}') = (R_n, \phi_n + 2\pi/3) - (R_m, \phi_m),$$

$$(R_{mn}'', \phi_{mn}'') = (R_n, \phi_n + 4\pi/3) - (R_m, \phi_m).$$

For the particular case of the interaction of the two triads $r(\mathbf{R}, \phi)_3$ and $r(\mathbf{R}, \phi + \pi/3)_3$, which together make up the hexad $r(\mathbf{R}, \phi)_6$, two hexads coalesce and we obtain

$$2r^2(\mathbf{R}, \phi)_6 + r^2(2\mathbf{R}, \phi)_6.$$

The complete S_2 for the hexad $r(\mathbf{R}, \phi)_6$ is therefore

$$6r^2(0, -) + 2r^2(\sqrt{3}\mathbf{R}, \phi + \pi/6)_6 + 2r^2(\mathbf{R}, \phi)_6 + r^2(2\mathbf{R}, \phi)_6.$$

These results make it possible to write down the S_2 for the most general trigonal distribution in S_1 , namely,

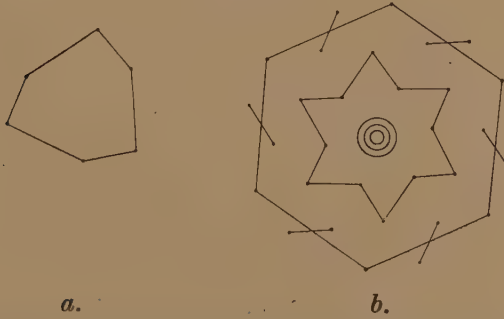
$$c(0, -) + \Sigma r_n(\mathbf{R}_n, \phi_n)_3.$$

It is

$$(c^2 + 3\Sigma r_n^2)(0, -) + \Sigma r_n^2(\sqrt{3}\mathbf{R}_n, \phi_n + \pi/6)_6 + c^2 \Sigma r_n(\mathbf{R}_n, \phi)_6 + \Sigma r_n r_m [(\mathbf{R}_{nm}, \phi_{nm})_6 + (\mathbf{R}_{nm}', \phi_{nm}')_6 + (\mathbf{R}_{nm}'', \phi_{nm}'')_6].$$

We now use these general results in the simple case (1) in which the

Fig. 2.



trigonal distribution comprises six points (and no intensity at the origin). The six points comprise two triads (fig. 2 a),

$$A_1 = r_1(\mathbf{R}_1, \phi_1)_3 + r_2(\mathbf{R}_2, \phi_2)_3,$$

and the S_2 distribution (fig. 2 b) is then

$$3(r_1^2 + r_2^2)(0, -) + r_1^2(\sqrt{3}\mathbf{R}_1, \phi + \pi/6)_6 + r_2^2(\sqrt{3}\mathbf{R}_2, \phi + \pi/6)_6 + r_1 r_1(\mathbf{R}_{12}, \phi_{12})_6 + r_1 r_2(\mathbf{R}_{12}', \phi_{12}')_6 + r_1 r_2(\mathbf{R}_{12}'', \phi_{12}'')_6,$$

namely, a point at the origin and five hexads.

We consider also the case (2) when these two triads lie at the corners of and midpoints of an equilateral triangle, namely (fig. 3 a),

$$A_2 = r_1(2\mathbf{R}/\sqrt{3}, \phi + \pi/6)_3 + r_2(\mathbf{R}/\sqrt{3}, \phi - \pi/6)_3,$$

in which case

$$V = 3(r_1^2 + r_2^2)(0, -) + r_1^2(2\mathbf{R}, \phi)_6 + (r_2^2 + 2r_1 r_2)(\mathbf{R}, \phi)_6 + r_1 r_2(\sqrt{3}\mathbf{R}, \phi + \pi/6)_6$$

and the five hexads of the case (1) coalesce to give only three hexads. The points which there gave two of the three hexads with intensity $r_1 r_2$ coalesce with the hexad of intensity r_2^2 to form a "compound" hexad of intensity $(r_2^2 + 2r_1 r_2)$.

Finally, we consider the case (3) when six points form a hexad, $r(R, \phi)_6$, and when, in addition, there is an intensity c at the origin (fig. 3 c). With this S_1 distribution

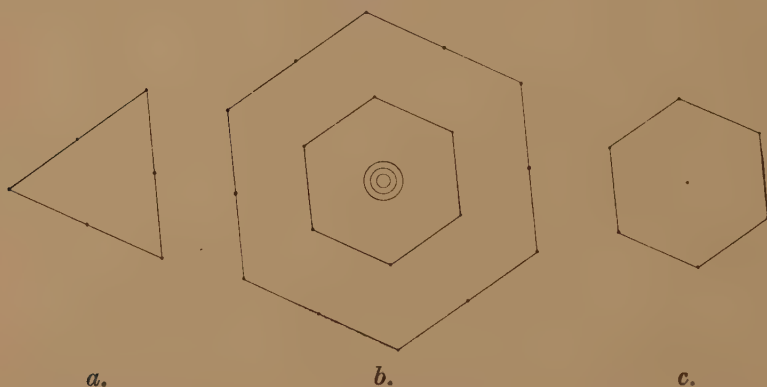
$$A_3 = c(0, -) + r(R, \phi)_6,$$

the S_2 distribution is

$$(c^2 + 6r^2)(0, -) + (2cr + 2r^2)(R, \phi)_6 + 2r^2(\sqrt{3}R, \phi + \pi/6)_6 + r^2(2R, \phi)_6.$$

It will be noticed that, when $c=0$, A_3 is a degenerate case of A_1 in which the two triads now form a hexad; correspondingly in S_2 two pairs of the

Fig. 3 a-c.



five hexads coalesce to give only three distinct hexads. The points, which in $V(A_1)$ form 2 hexads of intensity r_1^2, r_2^2 , now in $V(A_3)$ coalesce to give a single "compound" hexad of intensity $2r^2$. It should also be noticed that the point distribution $V(A_3)$ is identical with $V(A_2)$ shown in fig. 3 b whether or not $c=0$. The point intensity distributions differ in that $V(A_2)$ has two simple hexads and one hexad which is compound of order 3, whereas $V(A_3)$ has one simple hexad and two compound hexads of order 2 and 4 respectively.

4. *The Interpretation of Vector Maps.*

We have shown how to work out, in systematic fashion, the S_2 distribution corresponding to any trigonal S_1 distribution. The principles are the same for distributions with other symmetries. The companion problem as to the derivation of the interpretation of vector maps falls into two parts. (1) Given a point distribution V , what is the complete

class of S_1 point distributions which have this point distribution as their common vector map ? (2) If an intensity, positive, negative, or zero be assigned to each of the points which occur in V , what, in the case of each of the S_1 sets already found, is the complete set of intensities values which may be assigned to its points ? *

We indicate the method of attacking this problem by an example. Thus suppose that the vector point map † is given by (fig. 3 b),

$$V = (0, -) + (2R, \pi/6)_6 + (\sqrt{3}R, 0)_6 + (R, \pi/6)_6.$$

Evidently any corresponding S_1 point distribution is trigonal. It cannot consist of a single triad (S, ϕ) with or without a point at the origin, since this will not give three hexads.

Now let us take a general trigonal distribution

$$A = (0, -) + \Sigma(S_m, \phi_m)_3.$$

Among the terms in the corresponding V function are "domestic" terms, obtained from each individual triad, of the form

$$(\sqrt{3}S_m, \phi_m + \pi/6)_6 + (0, -).$$

Hence the only values of $(S_m, \phi_m)_3$ which are permissible are given by

$$\begin{aligned} S_m = R/\sqrt{3}, \quad \phi_m = 0 \text{ or } \pi/3: \quad S = 2R/\sqrt{3}, \quad \phi_m = 0 \text{ or } \pi/3: \\ S = R, \quad \phi_m = -\pi/6 \text{ or } +\pi/6. \end{aligned}$$

In fact, only six triads are satisfactory as regards their domestic contribution, namely,

$$\begin{aligned} (1) (R/\sqrt{3}, 0)_3, \quad (2) (2R/\sqrt{3}, 0)_3, \quad (3) (R, -\pi/6)_3, \\ (4) (R/\sqrt{3}, \pi/3)_3, \quad (5) (2R/\sqrt{3}, \pi/3)_3, \quad (6) (R, \pi/6)_3. \end{aligned}$$

If now we consider the interactions of these triads two at a time, we find that of the eighteen trios of vectors, one (or more) of each of the twelve sets of vectors 12, 13, 14, 16, 23, 25, 26, 34, 45, 46, 56 is inadmissible, and only the three pairs of triads 15, 24, 36 are satisfactory so far as this test is concerned. This means that there are at most three S_1 sets which give the required entries in S_2 , namely,

$$\begin{aligned} A_x = (0, -) + (R/\sqrt{3}, 0)_3 + (2R/\sqrt{3}, \pi/3)_3, \\ A_y = (0, -) + (R/\sqrt{3}, \pi/3)_3 + (2R/\sqrt{3}, 0)_3, \end{aligned}$$

* We assume at present that all points of the vector map are specified, including those points belonging to it which happen to have zero intensity. We do not here discuss the analysis of a vector point intensity map which, in addition to given entries with positive or negative intensities, has certain points on it which, happening to have the intensity zero, are not specified.

† Cf. Wrinch, J. Amer. Chem. Soc. lx. p. 2005 (1938).

(one of which is obtained by turning the other through two right angles about O_2) and

$$A_z = (0, -) + (R, -\pi/6)_3 + (R, \pi/6)_3.$$

One more test remains. The terms due to interactions of $(0, -)$ with each of these triads must give suitable hexads. But in A_x we have for the interactions of $(0, -)$ with each of the triads,

$$(R/\sqrt{3}, 0)_6 + (2R/\sqrt{3}, 0)_6,$$

both of which are inadmissible. Hence there is no $(0, -)$ point in A_x or A_y . On the other hand, the corresponding interactions for A_z give

$$(R, -\pi/6)_6 + (R, \pi/6)_6,$$

that is to say, two superposed hexads of suitable type. Hence the complete solution of the problem is that no point sets are possible other than (fig. 3 a)

$$A_2 = (R/\sqrt{3}, 0)_3 + (2R/\sqrt{3}, \pi/3)_3,$$

or

$$A_2' = (R/\sqrt{3}, \pi/3)_3 + (2R/\sqrt{3}, 0)_3,$$

or (fig. 3 c)

$$A_3 = (0, -) + (R, -\pi/6)_3 + (R, \pi/6)_3.$$

We now consider the same vector map with intensities attached, say

$$V = \mathbf{O}(0, -) + \mathbf{A}(2R, \pi/6)_6 + \mathbf{B}(\sqrt{3}R, 0)_6 + \mathbf{C}(R, \pi/6)_6.$$

If we now assign intensities to the points in A_3 , they must evidently be trigonally arranged, and we have, in general,

$$A_3 = c(0, -) + s(R, -\pi/6)_3 + t(R, \pi/6)_3,$$

whence

$$V = (c^2 + 3s^2 + 3t^2)(0, -) + st(2R, \pi/6)_6 + (s^2 + t^2)(\sqrt{3}R, 0)_6 \\ + (2st + cs + ct)(R, \pi/6)_6,$$

so that

$$c^2 + 3s^2 + 3t^2 = \mathbf{O},$$

$$st = \mathbf{A},$$

$$s^2 + t^2 = \mathbf{B},$$

$$c(s+t) + s^2 + t^2 = \mathbf{C}.$$

Evidently there is a possible set of intensities (c, s, t) only when

$$(\mathbf{C} - \mathbf{B})^2 = (\mathbf{O} - 3\mathbf{B})(\mathbf{B} + 2\mathbf{A}),$$

and when neither $(\mathbf{B} + 2\mathbf{A})$ nor $(\mathbf{B} - 2\mathbf{A})$ is negative. In this event, apart from an interchange of s and t , there are two and only two sets of intensities, namely, (c, s, t) and $(-c, -s, -t)$, where

$$c = (\mathbf{O} - 3\mathbf{B})^{\frac{1}{2}},$$

$$2s, 2t = (\mathbf{B} + 2\mathbf{A})^{\frac{1}{2}} \pm (\mathbf{B} - 2\mathbf{A})^{\frac{1}{2}}.$$

Correspondingly, in the case of the ground plan A_2 , we take the S_1 point intensity set as

$$A_2 = u(R/\sqrt{3}, 0)_3 + v(2R/\sqrt{3}, \pi/3)_3,$$

whence

$$V(A_2) = (3u^2 + 3v^2)(0, -) + v^2(2R, \pi/6)_6 + uv(\sqrt{3}R, 0)_6 \\ + (u^2 + 2uv)(R, \pi/6)_6.$$

Thus there is a possible pair of intensities (u, v) only when

$$3B^2 = A(O - 3A), \\ C = O/3 - A + 2B.$$

When these conditions are satisfied, there are two and only two pairs of intensities, namely, (u, v) and $(-u, -v)$, where

$$v = A^{\frac{1}{2}}, \quad u = (O/3 - A)^{\frac{1}{2}}.$$

As an example of a case in which the conditions are not satisfied and there is, consequently, no intensity distribution on the A_2 ground plan, we cite the case in which $B > A > C$.

5. Periodic Distributions and their Vector Maps.

So far we have been concerned only with finite distributions and their corresponding V functions. We consider now how to adapt these ideas to the case of distributions which are periodic.

We consider first the simplest of all periodic distributions, in which a point of intensity r is repeated at intervals a along a line. To find its vector map, we take successively 1, 2, 3, . . . n points, at intervals a along a line. The vector functions are

$$r^2(0), \\ 2r^2(0) + r^2(\pm a), \\ 3r^2(0) + 2r^2(\pm a) + r^2(\pm 2a), \\ \dots \dots \dots \\ nr^2(0) + (n-1)r^2(\pm a) + (n-2)r^2(\pm 2a) \dots + r^2(\pm \overline{n-1}a),$$

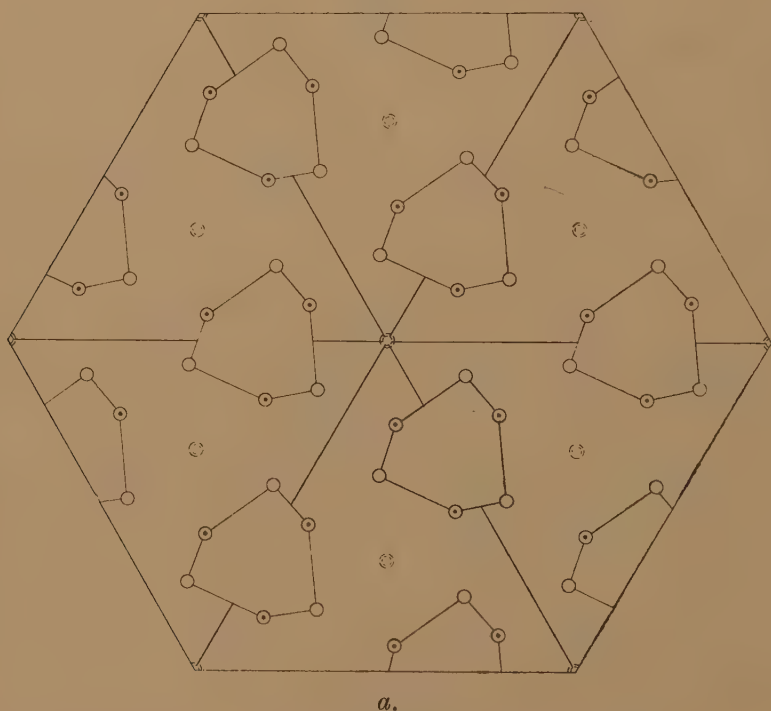
where $kr^2(\pm ma)$ means points of intensity kr^2 at $+ma$ and at $-ma$. To find the vector map of the periodic distribution, n must tend to infinity. But as n tends to infinity, entries in S_2 tend to infinity. In particular the largest entry (the entry at 0) is nr^2 when there are n points. If now all the intensities in S_2 are divided by the number of points, we obtain a function which tends to a finite limit as n tends to infinity. It comprises, in fact, the periodic distribution in which an entry r^2 occurs at O_2 (the origin in S_2) and at equal intervals a on either side of it.

This simple case indicates the convenient procedure to adopt. We take a finite distribution A and repeat it at intervals a along a line, *i. e.*, at the lattice points a , to give a periodic distribution, which may be called the a -lattice of A ,

$$L_a(A) = \sum_{k=-\infty}^{+\infty} A(x+ka).$$

We then build the vector map of $A(x)$ which, being centred at O_1 , may be called $V_A(O_1)$. Then the vector map of $A(x) + A(x+a)$ comprises $2V_A(O_1)$

Fig. 4.



together with $V_A(\pm a)$ (the distribution $V_A(0)$ repeated with $+a$ and $-a$ as centre). In general, we have

$$V\left(\sum_{k=0}^{n-1} A(x+ka)\right) = nV_A(O_1) + (n-1)V_A(\pm a) \\ + (n-2)V_A(\pm 2a) \dots + V_A(\pm \overline{n-1}a).$$

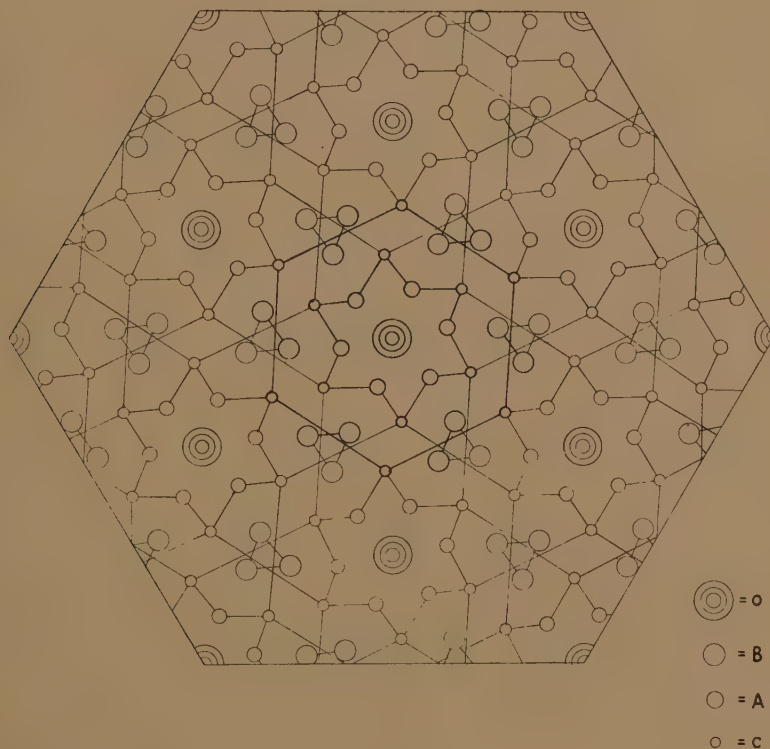
Now we define the vector function of the periodic distribution $L_a(A)$ by the equation

$$V(L_a(A)) = L_a(V_A(O_1)).$$

The procedure of building vector maps of periodic functions in S_1 is therefore as follows:—First find a finite distribution A in S_1 which repeated at all the a -lattice points gives the periodic function $L=L_a(A)$. Then build $V_A(O_1)$ in S_2 . Now repeat this distribution $V_A(O_1)$ at all the a -lattice points in S_2 and so form $L_a(V_A(O_1))$. Then $L_a(V_A(O_1))$ is $V(L_a(A))$, the vector map of L , i. e., of the a -lattice of A .

A periodic distribution L can be analysed into a repeating unit A in an

Fig. 4.



b.

infinite number of different ways. It is plain, however, that provided A and B are such that

$$L_a(A)=L_a(B)=L,$$

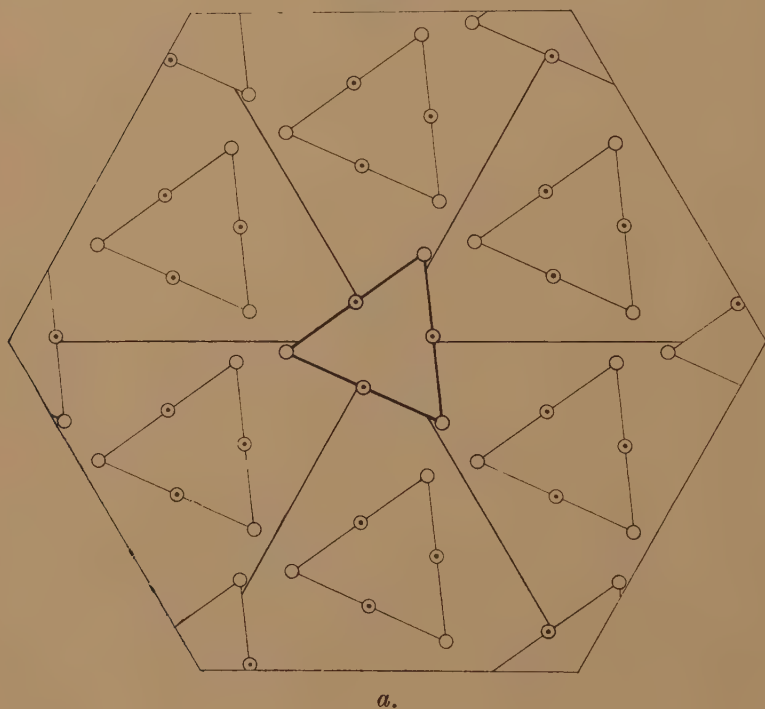
then

$$L_a(V_A(O_1))=L_a(V_B(O_1))=V(L).$$

In fact, whatever the unit A , provided only that $L_a(A)$ gives the required periodic distribution L_a , we can build $V(L_a)$ by building the a -lattice of $V_A(O_1)$. This argument, given for simplicity in terms of a one-dimensional lattice, applies unchanged to two and three-dimensional lattices.

As examples we take the units A_1' , consisting of A_1 (fig. 2 *a*) and a certain additional point, A_2 (fig. 3 *a*) and A_3 (fig. 3 *c*). Repeating each of them by parallel displacement at the points L of a certain rhombic lattice, we obtain three periodic distributions, $L(A_1')$, $L(A_2)$, $L(A_3)$, shown in figs. 4 *a*, 5 *a*, and 5 *c* respectively. To build the vector map of $L(A_1')$, we repeat the vector map of A_1' by parallel displacement at the same L points as shown in fig. 4 *b*. It has already been noticed that the point vector maps of A_2 and A_3 are identical; it follows that the periodic point vector

Fig. 5.



maps of $L(A_2)$ and $L(A_3)$ are identical, as shown in fig. 5 *b*. The relationship between the vector map of $L(A_1')$ and that common to $L(A_2)$ and $L(A_3)$, apparent from a comparison of figs. 4 *b* and 5 *b*, will be discussed later.

6. *Density Deviations in Periodic Distributions.*

It is of interest to notice the effect upon a vector function of certain changes in the intensities of the S_1 function from which it is derived, when it is periodic. Thus suppose the entry r_m at the point (ma) in an S_1

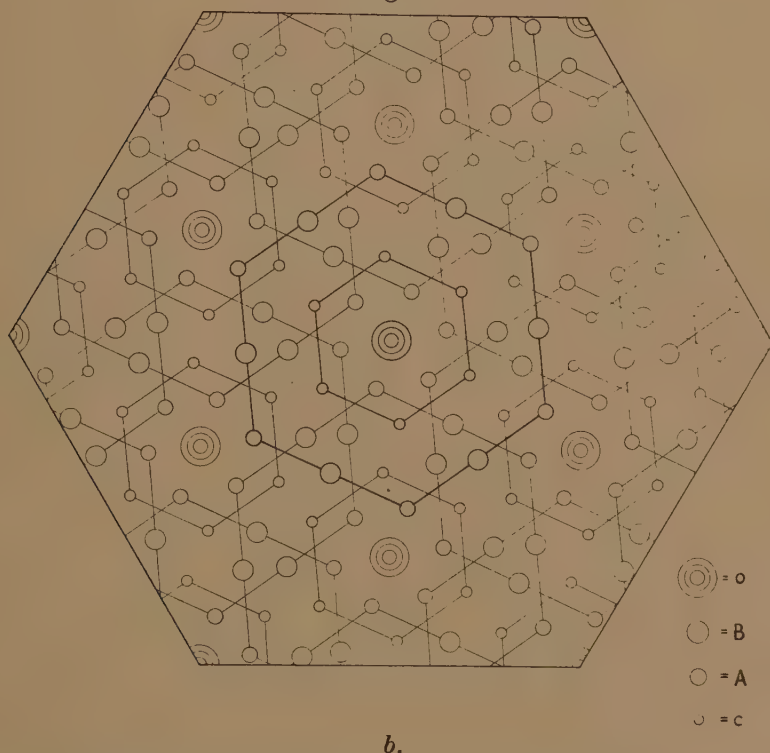
distribution A, where $r_{n+m}=r_m$ is changed to pr_m+q in a second S_1 distribution A'. Then whereas V_A has as its entries

$$\Sigma r_m^2(0), \quad \Sigma r_m r_{m+1}(\pm a), \dots$$

where the summation extends over n distinct terms in each case, the vector function $V_{A'}$ has as its entries similar summations of the form

$$\Sigma (pr_m+q)^2(0), \quad \Sigma (pr_m+q)(pr_{m+1}+q)(\pm a), \dots$$

Fig. 5.



But for a periodic intensity distribution, and only then,

$$\Sigma r_m = \Sigma r_{m+k} = n\bar{r},$$

where \bar{r} is the average value of r in the unit cell. Hence

$$\Sigma (pr_m+q)(pr_{m+k}+q) = p^2 \Sigma r_m r_{m+k} + nq^2 + 2npq\bar{r}.$$

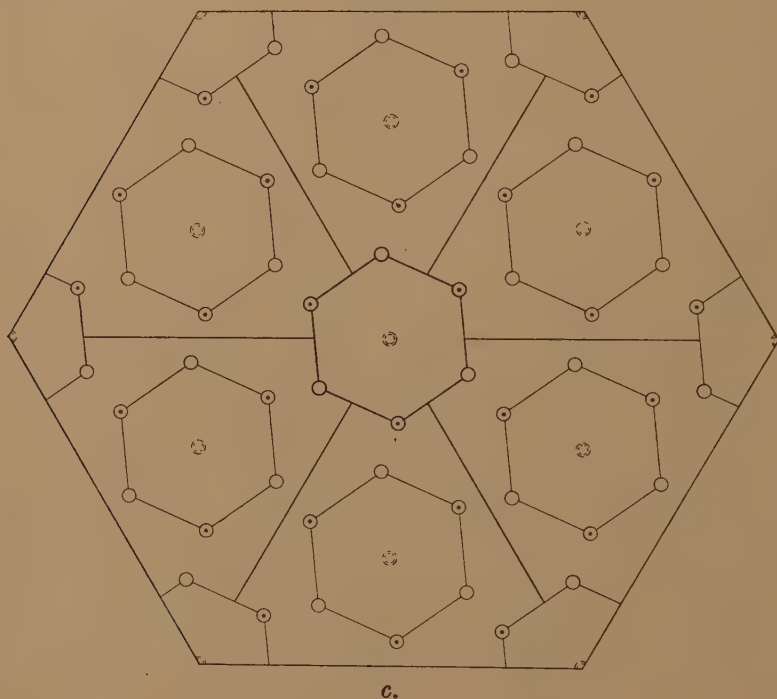
Thus at every point the entries in $V_{A'}$ are obtained by multiplying the entries in V_A by a constant and adding a second constant, a result which may be written in the form

$$V_{A'} = PV_A + Q,$$

where $P = p^2$, $Q = n(p\bar{r}+q)^2 - np^2\bar{r}^2$.

This result, proved for a one-dimensional periodic distribution, follows in a similar fashion for periodic distributions in two and three dimensions. It indicates that the geometry of a vector map in S_2 is independent of the scale on which the intensities in S_1 are measured, and of the zero from which they are measured. The set of points at which the vector function V_A takes a constant value V_0 still form a set at which the vector function $V_{A'}$ takes a constant value. The only difference is that the name attached to this contour is changed from V_0 to PV_0+Q , where P and Q are as

Fig. 5.



given above. Incidentally, it should be emphasized that vector maps (and therefore X-ray observations in general) do not distinguish between positive and negative intensity deviations. These results are of importance in practice in two distinct ways:—(1) It may be convenient in working out vector maps of complicated sets to deal, not with absolute intensities, but with deviations of intensity from some mean value. Then the result shows that it is necessary only to add a fixed amount to all vector map readings, which will then represent the absolute vector map. (2) On the other hand, suppose that, in the case of a given vector map, the intensities are measured from an unknown zero and on an

unknown scale, then the entries on the absolute vector map are certainly a linear function of the actual entries.

7. The Multiplicity of Units in a Periodic Distribution.

Let us first consider the multiplicity of units of a lattice L , which is defined by a unit 25 (two points of intensities 2 and 5 at unit distance apart) repeated at distances of 4 units along a line. This gives the 4-lattice function

$$L=L_4(25)=.25 \dots 25 \dots 25 \dots 25 \dots 25 \dots 25,$$

which may be written $|25 \dots 2|$. Now consider $l_4(25)$, the 4-lattice transformations of the unit (25), namely, all the units obtained by moving an entry in this unit by 4, 8, $\dots 4n$ steps, *i. e.*, to congruent points. These comprise the units $(5 \dots 2)$, $(5 \dots \dots 2)$ and so on. All these units yield again the same L_4 function. We remark that only two of the set of units $l_4(25)$ are non-overlapping in the lattice, namely (25) and $(5 \dots 2)$.

So far each entry in L has been allocated to one unit only. But consider the unit $(2 \dots 23)$ which also gives L as its 4-lattice function, though now one entry in L_4 is compound in the sense that more than one unit contributes to it. Consider, further, the 4-lattice transformations of this unit, in which any number of entries are moved to congruent points, all of which again give the original L .

If negative intensities are under consideration, further sets can be constructed, *e. g.*, $l_4(n.25\bar{n})$. Another example is the case already mentioned above, the lattice transformations of $(2 \dots 23)$ itself, also obtainable by superposing on $(\dots 25)$ the unit $(2 \dots \bar{2})$. Any such unit is obtained by arbitrarily adding to (25) any intensity at any point, and an equal and opposite intensity at a congruent point. Thus a unit with an arbitrarily large number of pairs of point intensities can be constructed.

It will be remarked that units with the minimum number of entries—which we will call “minimum” units—are obtained only when no entries in the lattice (zero or otherwise) are compound. Evidently it is in general *sufficient* to obtain for any lattice function L_4 the minimum units above, since all others are then derived by arbitrarily adding pairs of equal and opposite intensities at pairs of congruent points. However, it is not *necessary* to investigate even all the minimum units, since any one is a lattice transformation of any other.

The minimum units in this case are

$$l_4(25)=(25), (5 \dots 2), (5 \dots \dots 2), (5 \dots \dots \dots 2), \text{ etc.}$$

In the 4-lattice only the first two are non-overlapping, and of these, 25 having a “span” of 2 entries is more compact than $5 \dots 2$ having a “span” of 4 entries.

It seems convenient then in investigating any lattice to look for the minimum unit (or units) of minimum span. All other units can be derived from it (or them), if and when required.

From this discussion (which can be rewritten on exactly the same lines to deal with two or three-dimensional cases) we draw a number of conclusions. We seek, in interpreting a given S_2 distribution, to find one or more S_1 distributions which correspond to it. We would first emphasize that this type of analysis permits only the correlation of a periodic point distribution in S_2 , with a certain definite class of periodic point distributions in S_1 . Thus in applications to crystallography, it is plain that the idea of dividing up an S_1 crystal, *i. e.*, a periodic point distribution in S_1 in this or that way into molecules is entirely foreign to it. Nevertheless, particularly in the case in which there is known to be one molecule per cell in the crystal, it is useful to notice that the possibilities as to the structure of the molecule allowed by the derivation of each S_1 distribution can be systematically explored. Thus for each S_1 distribution, we investigate (1) the minimum unit (or units) of minimum span. Evidently these in turn may first be put forward as the structure of the molecule. If, and when data from other fields show these to be unacceptable, we can draw in regular order upon (2) other minimum units (obtained by shifting 1, 2 . . . entries to congruent points), (3) units obtained by adding to the previous units one, two . . . arbitrary points of equal and opposite intensities at congruent points. This discussion, in which we wish to emphasize the full extent to which there is a multiplicity of units all of which give the same lattice, shows that in practice the method of analysing S_1 distributions is clear, so that this multiplicity is no handicap to the systematic investigation of molecular structure (to be contrasted with the general crystal structure) in practice. It may be pointed out that there is in practice a severe limitation on the set of units which need be discussed. If we are dealing with a three-dimensional S_1 it is sufficient in dealing with minimum units to deal only with those which in the lattice are non-overlapping, if we are studying a crystal whose molecules are presumed to be non-interpenetrating. If we are dealing with a two-dimensional S_1 , which is the projection of a three-dimensional crystal, or with a one-dimensional S_1 , which is the projection of a three- or two-dimensional crystal, then some specifiable amount of overlapping may be permissible. Thus in the case cited above, only (25) and (5 . . . 2) of the minimum units need be discussed, if a one-dimensional crystal of non-interpenetrating molecules is under consideration; possibly one or even a few more may come into consideration if we are dealing with a projection of a two- or three-dimensional crystal.

Not only are there in practice limitations which reduce the units in category (2) to a very small number or zero, there are also limitations

in the extent to which units in category (3) have any appreciable degree of prior probability. For any unit in (3) is such that it contains one or more pairs of point intensities such that the intensities happen to be equal and opposite and the positions of the points happen to be congruent (a situation with a certain degree of improbability in practice in that the unit and the lattice then share a metrical parameter).

We now turn to periodic distributions in S_2 . The multiplicity of V units available for a given S_2 lattice may again be dealt with on the same lines. The complete analysis requires the consideration in turn of all the S_1 units corresponding to each of the units into which the periodic distribution can be analysed. Such an investigation, though it may be tedious in complicated cases, offers no difficulties of a theoretical nature. Thus consider the periodic S_2 distribution shown in figs. 4*b* and 5*b*. It consists of a point at O together with hexads, A , B , and C . We begin by considering the minimum unit

$$V = O + (A_1)_6 + (B_1)_4 + (C_1)_6,$$

namely, a point O together with the nearest A , B , and C hexads. This has no S_1 units corresponding and so is abandoned. A minimum unit which has corresponding S_1 units is obtained by taking (as shown in fig. 5*b*)

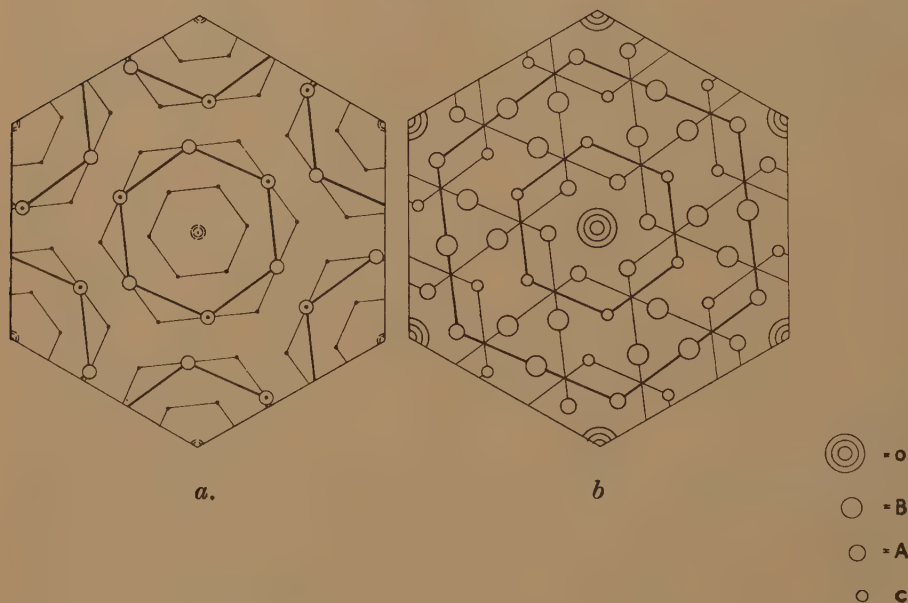
$$V = O + (A_2)_6 + (B_3)_6 + (C_1)_6,$$

namely, a point O together with the A hexad nearest but one, the B hexad nearest but two, and the nearest C hexad. To this correspond the units A_2, A_3 . Thus there correspond to the periodic S_2 distribution, the periodic atomic distributions shown in fig. 5*a* and in fig. 5*c*. There are no other minimum units. An example of a non-minimum unit is shown in fig. 4*b*, in which the unit comprises, in addition to a point O and the hexad A_1 , a share of the hexads B_1, B_2, C_1 , and C_2 . This unit corresponds to the unit A_1' in S_1 , and the analysis of the S_2 periodic distribution thus draws our attention to the S_1 periodic distribution shown in fig. 4*a*. This, it will be remarked, is identical with that shown in fig. 5*c*, since the unit (as has earlier been explained for a one-dimensional case) is simply a lattice transformation of the unit in fig. 5*c*. Thus nothing has been gained by considering this non-minimum unit in the S_2 distribution. In general this is the case, and a systematic investigation of the severely limited minimum units gives the complete solution of the problem.

Attention should, however, be directed to the special cases in which non-minimum units in S_2 can lead to S_1 distributions other than those obtained from the minimum units. One of these has already been discussed*.

* Langmuir and Wrinch, Proc. Phys. Soc. (in course of publication). Wrinch, Phil. Mag. xxvii. p. 98 (1939).

Suppose that a unit of appropriately degenerate type is built into a lattice expressly designed to ensure that certain inter-unit vectors make closed circuits with some intra-unit vectors. It can then happen that the addition of certain new points to the unit, while giving additional points in its vector map, yet does not lead to any new points in the lattice of the vector maps. This can only happen when the new points contributed by the vector map of a modified unit are superposed on points already contributed by the vector maps of other original units. Evidently such an S_1 is inevitably discovered from the periodic vector map when all the

Fig. 6 *a, b.*

V units into which the distribution can be analysed are considered in turn. This situation thus brings no practical difficulties in its train.

As an example consider the case shown in fig. 6 *b*, in which the individual vector maps shown in fig. 5 *b* are built into a lattice specially selected so that O , A , B , are collinear. The minimum unit

$$V = O + (A_2)_6 + (B_3)_6 + (C_1)_6$$

still gives a 7-point S_1 unit consisting of a hexad and central point, as in the case shown in fig. 5 *c*, or the 6-point S_1 unit consisting of points at the corners and mid-points of the sides of an equilateral triangle, as in the case shown in fig. 5 *a*. However, in this new case there are certain closed circuits of inter-unit and intra-unit vectors owing to the alignment of the

points **O**, **A**, **B**, and two new hexads can be added to the 7-point unit without any new points being added to the periodic S_2 distribution. The existence of this 19-point unit (shown in fig. 6 *a*) is at once detected when the non-minimum **V** unit consisting of **O** together with a share of the hexads **A**₁, **A**₂, **A**₃, **B**₁, **B**₂, **B**₃, **C**₁, **C**₂, **C**₃, **C**₄ comes in turn for consideration. It can evidently also be obtained by adding points to the 6-point unit shown in fig. 5 *a*.

XLV. *The Stress Distribution in an Aeolotropic Circular Disk.*

By H. ÔKUBO *.

[Received August 2, 1938.]

THE mathematical treatment of the elastic equilibrium of an aeolotropic substance in two dimensions was carried out by K. Wolf recently †, and its further study is developed by the writer ‡. Using the fundamental relation of the stress function in an aeolotropic substance proposed by K. Wolf, the writer will introduce the approximate solution of the stress components in an aeolotropic circular (or elliptic) disk subjected to two opposite forces in the direction of the axis of symmetry. And, so far as the writer is aware, no study to serve as a reference for the present subject is yet published. If we denote by χ the stress function in an aeolotropic substance, then χ satisfies the following equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2}{\partial x^2} + \frac{1}{k^2} \frac{\partial^2}{\partial y^2}\right)\chi = 0, \quad \dots \quad (1)\S$$

where $k^2 = E_1/E_2$, E_1 and E_2 denote Young's moduli in x , y directions respectively.

The equation (1) is readily transformed into

$$\left(\frac{\partial^2}{\partial x^2} + \frac{1}{k} \frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2}{\partial x^2} + \frac{1}{k} \frac{\partial^2}{\partial y^2}\right)\chi + \frac{(k-1)^2}{k^2} \frac{\partial^4 \chi}{\partial x^2 \partial y^2} = 0. \quad \dots \quad (2)$$

When k takes the value near unity (which is often the case for many crystals) the second term is neglected, being very small, and we have the equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{1}{k} \frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2}{\partial x^2} + \frac{1}{k} \frac{\partial^2}{\partial y^2}\right)\chi = 0.$$

Putting $\sqrt{k}y = y_1$, it becomes

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y_1^2}\right)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y_1^2}\right)\chi_1 = 0. \quad \dots \quad (3)$$

* Communicated by the Author.

† K. Wolf, *Z. f. Angew. Math. u. Mech.* H. 5, Bd. 15 (October 1934).

‡ H. Ôkubo, Science Reports of the Tôhoku Imperial University, series 1, xxv. no. 5 (March 1937).

§ K. Wolf, *loc. cit.*

The stress components are

$$X_x = k \frac{\partial^2 \chi_1}{\partial y_1^2}, \quad Y_y = \frac{\partial^2 \chi_1}{\partial x^2}, \quad X_y = -\sqrt{k} \frac{\partial^2 \chi_1}{\partial x \partial y_1}. \quad (4)$$

On the other hand, χ_1 is the stress function in an isotropic substance in x, y_1 plane, and if we denote by X'_x, Y'_y, X'_y the corresponding stress components, it becomes

$$X_x = kX'_x, \quad Y_y = Y'_y, \quad X_y = \sqrt{k}X'_y. \quad (5)^*$$

Hence the problem is reduced to finding the stress components in an isotropic substance in x, y_1 plane under the corresponding boundary conditions. When an aeolotropic circular disk of radius unity is compressed diametrically by two opposite forces in y direction, the problem is reduced to solving the problem under the elliptic boundary in an isotropic substance. (The ellipse has the major and minor axes as unity and \sqrt{k} respectively.) But such a case has been treated by the writer formerly †. Similarly, the problem of an aeolotropic square plate is reduced to that of an isotropic rectangular plate.

Numerical calculations :—

For an example, we will take a thin quartz plate which is cut out by

* Results given by (5) give the approximate value with sufficient accuracy when the difference of the elastic constants in x, y directions is comparatively small. For instance, if we take the case of the stress components in a semi-infinite domain having a boundary $y=0$, and compressed by a vertical force P operated at the origin, the stress components on y axis are

$$X_x=0, \quad Y_y = -\frac{k(1+k)}{\pi k^2 y} P, \quad X_y=0,$$

in an aeolotropic substance

$$X_x=0, \quad Y_y = -\frac{2P}{\pi\sqrt{ky}}, \quad X_y=0,$$

approximate values given by (5).

$$X_x=0, \quad Y_y = -\frac{2P}{\pi y}, \quad X_y=0,$$

in an isotropic substance (see H. Love, 'The Mathematical Theory of Elasticity,' 3rd Edition, p. 209).

If $E_1/E_2=1.30$, then $k=1.14$, the above equations give the value of Y_y as $-1.877 \frac{P}{\pi y}$, $-1.874 \frac{P}{\pi y}$, and $-2 \frac{P}{\pi y}$ respectively.

† H. Ôkubo, Proc. Phys. Math. Soc. Japan, 3rd ser. xvi. (March 1934).

a plane containing the principal axis and vertical to the plane of symmetry. In this case the strain-energy-function is

$$W = \frac{1}{2}c_{11}e_{xx}^2 + \frac{1}{2}c_{33}e_{yy}^2 + \frac{1}{2}c_{44}e_{xy}^2 + c_{13}e_{xx}e_{yy},$$

where $c_{11}=868$, $c_{33}=1074$, $c_{44}=582$, $c_{13}=134$ *.

Since Young's moduli in x and y directions are $c_{11}-c_{13}^2/c_{33}$ and $c_{33}-c_{13}^2/c_{11}$, we have

$$k = \sqrt{\frac{c_{11}}{c_{33}}} = 0.899.$$

Now, consider the following transformation,

$$x + iy = c \cosh w,$$

where $y_1 = \sqrt{ky}$, $w = \alpha + i\beta$.

Then we have

$$\left. \begin{aligned} x &= c \cosh \alpha \cos \beta, \\ y_1 &= c \sinh \alpha \sin \beta. \end{aligned} \right\} \quad \dots \dots \dots (6)$$

Equation (6) shows that the condition $\alpha = a$ represents the circumference of an ellipse having the major and minor axes unity and \sqrt{k} respectively, provided $c = 0.3177$ and $e^a = 6.131$.

When an elliptic plate is compressed by a pair of opposite forces P acting at the end of the minor axis, the stress components are

$$\left. \begin{aligned} \widehat{\alpha\alpha} - \widehat{\beta\beta} - 2i\widehat{\alpha\beta} &= \frac{1}{c \sinh(\alpha - i\beta)} \{ic \sinh \alpha \sin \beta \phi_1'(w) + \phi_2(w)\}, \\ \widehat{\alpha\alpha} + \widehat{\beta\beta} &= \text{real part of } [\phi_1(w)], \end{aligned} \right\} \quad (7)$$

where

$$\phi_1(w) = \sum_{n=0}^{\infty} A_n \cosh 2nw, \quad \phi_2(w) = c \sum_{n=0}^{\infty} B_n \sinh (2n+1)w,$$

and

$$\begin{aligned} A_0 &= \frac{2P}{\pi} \sum_{k=0}^{\infty} (-1)^{k+1} \frac{\sinh 2ka + \sinh 2(k+1)a}{k \sinh 2a + \cosh 2ka \sinh 2(k+1)a}, \\ A_n &= \frac{4P}{\pi} \sum_{k=n}^{\infty} (-1)^{k+1} \frac{\sinh 2ka + \sinh 2(k+1)a}{k \sinh 2a + \cosh 2ka \sinh 2(k+1)a}; \quad n \geq 1 \\ B_0 &= -A_0 + \frac{1}{2} \cosh 2a \cdot A_1, \\ B_n &= -\frac{(2n+1) \sinh a}{2 \sinh (2n+1)a} \{ \cosh 2na \cdot A_n - \cosh 2(n+1)a \cdot A_{n+1} \}, \quad n \geq 1 \end{aligned}$$

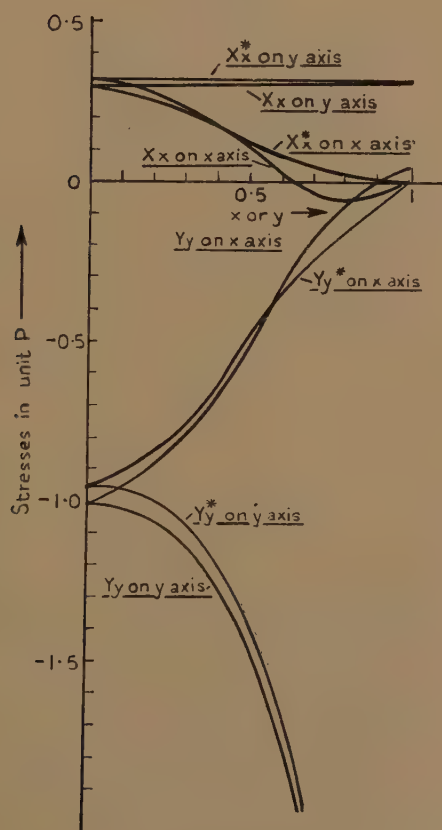
* Values of elastic constants used here are due to W. Voigt.

TABLE (unit P).

$x.$	$y.$	$\alpha\alpha.$	$\beta\beta.$	$X_x.$	$Y_y.$	$X_x^*.$	$Y_y^*.$
0	1.00	$-\infty$	0.351	0.316	$-\infty$	0.318	$-\infty$
0	0.75	-2.739	0.339	0.305	-2.739	0.318	-2.592
0	0.50	-1.455	0.335	0.301	-1.455	0.318	-1.379
0	0.25	-1.097	0.335	0.301	-1.097	0.318	-1.040
0	0	-1.008	0.335	0.301	-1.008	0.318	-0.955
0.25	0	-0.839	0.253	0.228	-0.839	0.248	-0.809
0.50	0	0.117	-0.515	0.105	-0.515	0.115	-0.497
0.75	0	-0.067	-0.110	-0.060	-0.110	0.025	-0.202
1.00	0	0	0.032	0	0.032	0	0

Where * shows values in an isotropic circular disk.

Fig. 1.



Values of constants A_n and B_n calculated from the above equation are (unit P) :—

$$\begin{array}{ll}
 A_0 = -0.6031, & B_0 = 1.2393, \\
 A_1 = 0.6763 \times 10^{-1}, & B_1 = -0.9885 \times 10^{-1}, \\
 A_2 = -0.1803 \times 10^{-2}, & B_2 = 0.4384 \times 10^{-2}, \\
 A_3 = 0.4793 \times 10^{-4}, & B_3 = -0.1633 \times 10^{-3}, \\
 \vdots & \vdots \\
 A_n = 95.76 \times (-1)^{n+1} e^{-2(n+1)a}, & B_n = 285.8 \times (-1)^n (2n+1) e^{-(2n+3)a}, \\
 \vdots & \vdots
 \end{array}$$

Values of stress components on the horizontal and vertical diameters calculated from the above equations are tabulated in the table on p. 511, and to compare the results with those of an isotropic one, values of stress components in an isotropic circular disk under the same boundary conditions are also added in the table. They are plotted in the figure.

XLVI. *Notices respecting New Books.*

Ultrasonics and their Scientific and Technical Applications. By Dr. LUDWIG BERGMANN, translated by Dr. H. STAFFORD HATFIELD. [Pp. 264+ix.] (London : G. Bell and Sons, Ltd. Price 16s. net.)

ALL who are interested in this subject, whether from the research standpoint or from the teaching standpoint, will heartily welcome this translation of Professor Bergmann's book, "Ultraschall." The translation is clear and reads smoothly and the phraseology is characteristically English rather than German. The book is essentially experimental in character and points of theoretical importance are discussed in general terms without recourse to detailed mathematical analysis. This does not imply that the theoretical side of the subject is vaguely or inadequately treated; on the contrary, the author has set out very lucidly the physical principles involved in such a way that although little mathematical knowledge is required a sound knowledge of many physical, especially optical, ideas is necessary. The reader may find it necessary, for example, to brush up his knowledge of the theory of optical diffraction gratings before he will appreciate fully Professor Bergmann's discussion of the optical methods of studying ultrasonic waves.

The scope of the book can only be briefly indicated. Pp. 1-38 deal with the generation of ultrasonic waves; pp. 39-103, their detection and measurement; pp. 104-160, their velocity and absorption in liquids and gases; pp. 161-188, their velocity in solids and the determination of elastic constants; pp. 189-226, further applications of ultrasonics. It is impossible to pick out any one section of the book as being of particular interest, for this will depend on the individual reader; the first and last sections will probably appeal most to those interested in the application of ultrasonics. Others will find those sections dealing with the interaction of light waves and ultrasonic waves particularly fascinating; a very wide range of diffraction phenomena is obtained and one's interest in this aspect of the subject is enhanced by the many beautiful photographs which are reproduced.

A notable feature of the book is the large number of references collected in the bibliography; 483 were given in the original German edition and almost another 100 are added in the present edition. The titles of all the papers are given in full, with the result that the bibliography occupies a considerable fraction of the whole book, actually 32 pages out of a total of 264 pages, *i. e.*, practically one eighth of the book. To the reviewer this seems excessive. It is very questionable whether any useful purpose is served by giving the full titles of original papers; these require generally two or three lines of print and occasionally four and even five lines. One wonders whether the references could not have been given in small print at the foot of the page in the usual way and the titles omitted; a consecutive numbering of the references throughout the book could still have been retained.

There appears to be a number of minor slips in the text; for example, on p. 75 there appears "grating consonant"; on p. 80, "travelling to the Z direction"; the footnote on p. 80 refers the reader to p. 74 of R. W. Wood's 'Physical

Optics,' p. 84 is probably intended ; on p. 86, reference (322) should be reference (332) ; on p. 87, we are referred to a trough T in figure 77, but the letter T does not appear ; on p. 89, we are told that a certain method was described " at the end of 1932," but the references quoted are dated 1934 ; on p. 117, we read " this phenomena." These are mainly trivial comments, but the reviewer wonders how many more he could find by diligently searching through the whole book.

To sum up : the book as a whole is good and the reviewer enthusiastically recommends it ; the bibliography occupies too much space and a less expensive book might have been produced by giving the references in a more concise form ; the text contains a number of minor slips. G. W. B.

The Principles of Statistical Mechanics. By RICHARD C. TOLMAN. [Pp. xix+661.] (Oxford : The Clarendon Press, 1938. Price 40s.)

THIS massive book is entirely different, in spirit and outlook, from the earlier one by Prof. Tolman with a similar title. Its somewhat repellent bulk is in part accounted for by the fact that in addition to a treatise on statistical mechanics, it incorporates a treatise, and a very excellent one, on the principles of quantum mechanics and another, perhaps rather an epitome than a textbook, on classical dynamics.

The chapters on statistical mechanics proper are themselves largely duplicated, the first half of the book setting out the treatment on Newtonian mechanical principles, and the second half not merely extending, but rebuilding from the beginning on the basis of the quantum theory. As a result, the reader's attention is directed throughout to underlying principles rather than to specific results and, moreover, the differences in outlook between the older and the newer methods are brought into prominence, because a strictly parallel treatment is not possible at all points, and the departures from it catch the attention at once.

In the section on the classical statistical mechanics, Boltzmann's H-theorem must always be a central (or perhaps a culminating) point, and here the writer has given an admirable summary of the meaning of the theorem and the conditions under which it holds, though it is almost needless to say that his argument is based on that of Gibbs and not that of Boltzmann. In an earlier section, it is interesting to note that his views on the ergodic hypothesis differ appreciably from those which he expressed in his earlier work, and that they have, probably unconsciously, been affected by the analysis of the nature of physical argument which was one of the features of the foundation period of the quantum mechanics.

The most immediately valuable part of the book is that in which modern statistical mechanics is built up. The two reasons for taking averages—the one, truly statistical, due to our postulated imperfection of knowledge of the system under consideration, and the other, due to the probability nature of the mechanics itself, are clearly explained. Then the analogue to the ergodic hypothesis is introduced and fully discussed, and the three common ensembles are introduced. The complications due to symmetry requirements unknown in classical mechanics, and to spin, are explained at length.

The machinery having been thus set up, a chapter is devoted to the Maxwell-Boltzmann, Einstein-Bose and Fermi-Dirac distributions, and another to the change in quantum mechanical systems with time. Then follows a very detailed discussion of the quantum analogue of the H-theorem. The third part of the

book is entitled "Statistical Mechanics and Thermodynamics" and, like the first two, deals rather with principles than with detailed applications. For these, the author refers the reader to Fowler's 'Statistical Mechanics,' which he considers will meet the needs of the manifold applications of statistical mechanics to physics and chemistry for a long time. It seems at least as certain that the book now under review will for as long a time continue to meet the needs of those who are not satisfied to know how to work the machine, but ask how and why it works.

J. H. AWBERRY.

A Preface to Mathematics. By C. E. VAN HORN. [Pp. xii+124.] (Chapman & Grimes, Inc., Boston, 1938.)

THE aim of Professor Horn is the admirable one of enlarging the outlook of prospective and actual teachers of Mathematics by providing material and viewpoints which are not given by their usual text-books. There is much in this little book which will contribute to this end, but it would possibly have been achieved more readily if the style were relieved by lighter touches, and if less routine mathematical manipulation had been included.

The subjects discussed range from the schoolboy "proof" that $1=2$, to the theory of the Riemann Integral and the convergence of infinite integrals. In each chapter there is a great deal to make the student think, and think profitably, about the theories and theorems which he has absorbed, but probably not digested, during his undergraduate studies. The effect will be improved in future editions, however, if the too numerous misprints are corrected.

Theory of Equations. By J. M. THOMAS. [Pp. x+211.] (McGraw Hill Publishing Company, Ltd., 1938. Price 12s.)

THIS book contains the elementary theory of equations together with chapters which are preliminary to a proposed later book in which the more advanced Galois theory will be discussed. It is intended primarily for graduate and advanced undergraduate students (see the preface), but most of it would be covered in the first year of an English honours course in Mathematics. In either case, a chapter on the Argand diagram, with a rather casual proof of De Moivre's Theorem, and another on determinants, seem superfluous.

There is a good account of systems of linear equations, symmetric functions, the sequences of Fourier-Budan and Sturm, and the solution of the cubic and quartic, but there is no mention of the general Tschirnhausen transformation. There is an excellent chapter on ruler and compass constructions, leading to the results of Gauss on regular polygons which can be constructed in this way, and the book ends with an introduction to the theory of simultaneous systems of non-linear equations.

There is some looseness in the terminology used. The term *imaginary* number frequently occurs, when *complex* number is intended, and there seems no reason for replacing the usual name *cofactor* by *algebraic complement*. On the whole, however, the book is an adequate introduction to the theory.

Modern Higher Algebra. By A. ADRIAN ALBERT. [Pp. xiv+319.] (Cambridge University Press, 1938. Price 18s.)

MODERN algebra is based upon the three concepts *group*, *ring*, and *field*, but until now there has been no book in English which develops the subject

systematically from these ideas. Professor Albert has therefore filled an obvious gap in mathematical literature by writing this book, which provides an account of algebra comparable with that given in van der Waerden's 'Moderne Algebra.' Certain subjects dealt with in the latter work are, however, reserved by the American writer for treatment in a later volume.

After a preliminary account of groups and rings, the theory of matrices is developed in considerable detail, including the theory of equivalence and canonical forms. Finite groups are next discussed, and algebraic fields over a given field—i. e., fields obtained by adjunction (to the given field) of the roots of an equation. This naturally leads to the Galois theory of equations, completed by the proof that the general equation of degree greater than four is insoluble by radicals. The remaining chapters give an introduction to the theory of algebras, particularly the algebras of matrices, and the theory of ideals.

The book provides a lucid and most acceptable addition to the few available works on modern algebra. It is well printed, and its value is enhanced by a considerable number of well-chosen examples.

The George Fisher Baker Non-Resident Lectureship in Chemistry at Cornell University. Chemical Kinetics. By Professor FARRINGTON DANIELS. [Pp. viii and 273.] (Cornell University Press, 1938. Price 15s.)

THE George Fisher Baker Non-Resident Lecturer at Cornell University is expected to present the most recent advances and the results of his own investigations in the field selected for discussion. It is particularly fortunate, therefore, that the lectures on Chemical Kinetics should have been delivered by Professor Farrington Daniels, whose well-known researches have penetrated into so many branches of this fascinating subject. In spite of the wide range of his work, however, Professor Daniels has made no attempt to cover the whole field, and those readers who expect to find all the main aspects of Chemical Kinetics treated in their proper perspective will be disappointed. The book opens with a discussion of general principles, followed by a more detailed treatment of unimolecular reactions. Bimolecular reactions, complex reactions, and chain reactions, are only mentioned briefly. The account of solution reactions, photochemistry, electrical activation, infra-red spectroscopy, the theoretical calculation of activation energies, and isotropic tracers, with which the rest of the book is occupied, is in no case entirely general, and one feels that sometimes space has been given to the description of relatively unimportant reactions. On the other hand, it has been possible in this way to link theory and experiment in an unusually successful fashion. Moreover, the author's attractive style and his lucid exposition of some of the most recent developments in chemistry, with which his own boldly conceived and carefully executed experiments are intimately connected, make the book one which can be read with profit by all interested in chemistry and by research students in particular.

R. SPENCE.

[The Editors do not hold themselves responsible for the views expressed by their correspondents.]

XLVII. *The Diffraction of Wireless Waves round the Earth. (A Summary of the Diffraction Analysis, with a Comparison between the Various Methods.)*

By G. MILLINGTON, Research and Development Department,
Marconi's Wireless Telegraph Co., Ltd., Chelmsford, Essex *.

[Received November 17, 1938.]

1. *Introduction.*

THE problem of the diffraction of wireless waves round the earth, taking account of the earth constants and of the height of the transmitter and receiver above the earth, has been attacked during the last two or three years by three independent methods. Wwedensky ⁽¹⁾ showed that Watson's original analysis, which was formally complete, could be extended in detail to cover the more general cases with which one is in practice concerned. van der Pol and Bremmer ⁽²⁾ also extended Watson's technique of converting the harmonic series into a contour integral, and by making a very detailed study of the intricate Hankel functions involved, they obtained a more accurate solution of the problem.

In addition, van der Pol and Bremmer developed the harmonic series in a totally different way by splitting it up into an infinite number of other series, each of which corresponds to a ray in the Airy treatment of the problem of the rainbow. In this way they have obtained a formula for an elevated transmitter and receiver, when the distance between them is small, in terms of a direct ray and a reflected ray, which allows for the divergence on reflexion at a curved surface, and contains a modified reflexion coefficient appropriate to the curved earth case. They point out that the modification produced to the simple ray treatment assuming reflexion at a flat surface is not large. For distances beyond the limit at which this simple reflexion formula is applicable, there is a region where the series derived from the contour integral is unwieldy. But by summing the series in several particular cases to as many as twenty terms, they have been able to verify that their two different ways of developing the harmonic series are really equivalent.

The third method is the application of the phase integral theory to the problem, as was done by T. L. Eckersley and the present writer

* Communicated by T. L. Eckersley, F.R.S.

in a recent paper ⁽³⁾. Where the appropriate boundary conditions are applied, this method leads to a complete solution of the problem ; but its special merit lies in the fact that it extends Watson's original solution in a relatively simple way, thus providing an alternative to the extremely complex analysis otherwise involved. It gives a clearer picture of the underlying physical principles, and it is also more general in that it shows the form of the further modification resulting from the inclusion of a gradient of refractive index in the atmosphere immediately above the earth.

As the notation which has been used in the development of these three methods is so different, and the final results are expressed in such dissimilar forms, it is difficult without a close reading of all the papers to see whether they are really equivalent. But it is obviously a matter of some importance to compare them, and the purpose of this paper is to show that the differences involved are only a question of the approximations made in deriving the results, and that with one serious exception they are not of any practical significance.

2. The General Form of the Complete Solution.

In discussing the analysis we will keep as far as possible to the notation used in our previous paper, but for convenience we will define the symbols again as they occur. Watson's original solution, for both transmitter and receiver on the surface of a perfectly conducting earth, may be put into the form

$$E = \sum_{s=0}^{s=\infty} A_s \exp[i(\omega t - n_s \phi)] \quad . \quad . \quad . \quad . \quad . \quad (2.1)$$

$|E|$ is the field strength due to a vertical dipole, and ϕ is the angular distance of the receiver from the transmitter, and n_s is the proper value which is determined from the proper value equation in the form

$$n_s = x + \rho_s x^{1/3} e^{-i\pi/3} \quad . \quad . \quad . \quad . \quad . \quad (2.2)$$

x is here given by

$$x = \frac{2\pi r_0}{\lambda}, \quad . \quad . \quad . \quad . \quad . \quad (2.3)$$

where r_0 is the radius of the earth, and λ the wave-length. ρ_s is a real constant, which for the first term, when $s=0$, is 0.8083.

It should be noticed that we use the time factor $\exp(i\omega t)$, whereas Wwedensky and van der Pol and Bremmer use the factor $\exp(-i\omega t)$, and in comparing their results with ours we therefore reverse all the phases. On substituting (2.2) in (2.1) we obtain a term $\exp[i(\omega t - x\phi)]$ which represents a progressive wave. As it is common to all the terms

of the series, we can omit it in considering the value of $|E|$. The remaining part of the exponential term in (2.1) contains an amplitude and a phase term, and we may write (2.1) as

$$E = \sum_{s=0}^{s=\infty} A_s \exp \left[\frac{-\beta_s \phi}{\lambda^{1/3}} \right] \exp \left[-i \frac{\beta_s \phi}{\lambda^{1/3}} \cot \frac{\pi}{3} \right], \quad \dots \quad (2.4)$$

where

$$\beta_s = (2\pi r_0)^{1/3} \rho_s \sin \frac{\pi}{3}. \quad \dots \quad (2.5)$$

The amplitude term A_s depends upon the distance d of the receiver from the transmitter, and on the wave-length. If we measure d and λ both in km., and express $|E|$ in mv./m. for 1 kW. radiated, the absolute value of A_s is

$$A_s = \frac{27.5}{\rho_s d^{1/2} \lambda^{1/6}} \quad \dots \quad (2.6)$$

where we have again omitted phases common to all the terms.

Now all three methods of extending Watson's solution to the general case agree in presenting the final result as a number of modifications imposed upon (2.4). These may be classed under four headings:—

- (a) The modification to ρ_s .
- (b) The corresponding modification to β_s .
- (c) The modification to A_s .
- (d) The inclusion of a height-gain factor F_s .

These modifications take the following forms:—

(a) For a given value of s , ρ_s is dependent on the earth constants, and is obtained as the solution of a generalized equation. The modified value of ρ_s is now complex, and may be written

$$\rho_s = |\rho_s| \exp(i\omega),$$

where the phase angle ω is, of course, different from ω used in the time factor.

(b) The exponential attenuation factor is modified simply by the use of this generalized value of ρ_s , so that β_s in (2.5) becomes

$$\beta_s = (2\pi r_0)^{1/3} |\rho_s| \sin \left(\frac{\pi}{3} - \omega \right). \quad \dots \quad (2.7)$$

(c) In the value for A_s in (2.6) we must now use the modified form of ρ_s , but there is also another factor introduced, which we have written as C in the denominator. Thus we now have

$$A_s = \frac{27.5}{C \rho_s d^{1/2} \lambda^{1/6}}, \quad \dots \quad (2.8)$$

where C is complex and is a function of ρ_s and the earth constants.

For any given values of ϵ and σ , ρ_s passes from the upper limit to the lower limit as λ is increased, there being a steep transition region in which most of the change occurs. van der Pol and Bremmer have given a set of curves from which the transition curve for any chosen values of ϵ and σ can be evaluated. But, as we have pointed out, under all likely practical conditions, when ρ_s has departed appreciably from the upper limit, $2\sigma\lambda c$ has become $\gg \epsilon$, and thus in practice a single transition curve can be drawn as a function of $\sigma^{1/2}\lambda^{5/6}$. In our paper we gave such curves for $|\rho_s|$ and ω for $s=0$, and for β_s for $s=0$ and 1.

In the next section we discuss the comparison of the tangent and Hankel approximations for β_s , but, as an additional check on the general equivalence of the two approximations, we have computed the curve corresponding to our curve for the phase angle ω from the curves for the real and imaginary parts of τ_s given by van der Pol and Bremmer, where their quantity τ_s is equivalent, except for a constant phase $\pi/3$, to our quantity ρ_s . Our curve for ω agrees well with this derived curve, the peak occurring in the same place, and having a maximum value of 20° for $s=0$ compared with their 22° .

4. The Exponential Attenuation Factor.

The quantity ρ_s has no dimensions, but from (2.7) we see that β_s has the dimensions of $r_0^{1/3}$. Following Watson's original choice of units we express the wave-length in (2.3) in km., and writing $r_0=6366$ km. we get $(2\pi r_0)^{1/3}=34.2$. Thus we can write (2.7) as

$$\beta_s = 34.2 |\rho_s| \cdot \sin\left(\frac{\pi}{3} - \omega\right),$$

and by putting $\rho_s=0.8083$ and $\omega=0$, we get Watson's lower limit of 23.94 .

Now van der Pol and Bremmer point out that their transition curve for β_s , using the Hankel approximation, agrees at the lower limit with this accurate Watson value, whereas the tangent approximation gives a value of 26.2 ; and for the upper limit the latter gives a value of 54.7 compared with the more accurate Hankel value of 55.0 . The curves thus agree very well at the upper limit, and they also keep close together down the steep part of the transition region, but they diverge markedly towards the lower limit.

The tangent approximation, as derived both from the Debye approximations to the Hankel functions as used by Wwedensky, and from the phase integral *, introduces an angle $\pi/4$, and in our paper we called this angle $q\pi$. We then adopted the device of altering q slightly from the

* See Appendix.

value of $\frac{1}{4}$ so as to adjust the lower limit of ρ_s to equal Watson's value. We pointed out that this device in no way affects the general form of the analysis, since the fundamental results are all expressed in terms of $q\pi$ and are independent of the precise value allotted to q . For purposes of continuity we retained our modified value of q throughout the transition curve. Comparison with the Hankel curve shows that the resulting curve now fits closely at the lower end of the curve and well up the steep transition region, but the upper limit for β_s is now 53.0.

It thus appears that the unmodified tangent approximation with $q=\frac{1}{4}$ is closer to the Hankel approximation at the upper end of the curve. Actually by means of the same device we could adjust the tangent curve to have an upper limit of 55.0, and by computing the upper end of the curve with q so adjusted, and the lower end of the curve with the existing adjustment, the two parts of the curves would meet with very little discontinuity in the middle of the transition region.

5. The Amplitude Factor.

As stated above, all the methods agree in representing the modification to the amplitude factor A_s as a term C in the denominator, in addition to the modification to ρ_s itself. We have given C in the form

$$C = 1 - \frac{1}{\zeta^2 \alpha_0^2} - \frac{\epsilon}{\zeta x \alpha_0^4}, \quad \dots \dots \dots (5.1)$$

where

$$\alpha_0^2 = -2\rho_s x^{-2/3} e^{-i\pi/3} \dots \dots \dots (5.2)$$

By writing

$$\eta = \frac{x^{1/3} e^{i\pi/6}}{\zeta \sqrt{2\rho_s}} \dots \dots \dots (5.3)$$

(5.1) may be written

$$C = 1 + \eta^2 + \frac{\eta}{(2\rho_s)^{3/2}} \dots \dots \dots (5.4)$$

As Wwedensky uses the tangent approximation, his modification to the amplitude factor is essentially equivalent to ours. He expresses his result, however, in a slightly different form. In effect he writes (2.8) as

$$A_s = \frac{0.047}{a_s \alpha^{1/2} \lambda^{1/6}} \dots \dots \dots (5.5)$$

where

$$a_s = 2(2\pi r_0)^{-2/3} \rho_s C = 0.00171 \rho_s C \dots \dots \dots (5.6)$$

He thus includes the factor ρ_s in his term a_s , whereas we have left it as a separate factor to show more directly the replacement of Watson's lower limit value by the generalized form. The inclusion of ρ_s modifies the shape of his curve for $|a_s|$ as compared with our curve for $|C|$

as functions of $\sigma^{1/2}\lambda^{5/6}$. The marked drop which we found after a slight initial increase with decreasing values of $\sigma^{1/2}\lambda^{5/6}$ is removed, and the curve for $|a_s|$ increases gradually at first and then more rapidly, until it shows the same steep increase that is eventually exhibited by the curve for $|C|$.

van der Pol and Bremmer have shown that the use of the Hankel approximation gives the same form for C , except that the third term in (5.4) must be omitted. For long waves, or for perfect conductivity, when $\eta \rightarrow 0$, $C \rightarrow 1$, and for short waves, when $\eta \gg 1$, the second term predominates. Computation shows that the omission of the third term only produces a small modification to the curve we have given for $|C|$. The slight initial increase that we found is removed, while the drop below the value of unity is somewhat increased, the minimum becoming 0.83 compared with our value of 0.95.

When the second term of C predominates we have

$$C = \eta^2 = \frac{(2\pi r_0)^{2/3} \lambda^{-2/3} e^{i\pi/3}}{\zeta^2 \cdot 2\rho_s}.$$

Thus from (5.6)

$$a_s = \frac{\lambda^{-2/3} e^{i\pi/3}}{\zeta^2},$$

and from (5.5)

$$A_s = \frac{0.047 \zeta^2 \lambda^{1/2} e^{-i\pi/3}}{d^{1/2}}. \quad \dots \quad (5.7)$$

In computing $|E|$ for a case in which several terms of the diffraction formula have to be taken, the phase term $-\pi/3$ can be omitted, unless we have to include any terms for which the second term in C does not predominate.

6. Comparison of Ground Curves.

Before considering the height-gain factor F_s , we can compare the three methods as far as the computation of ground curves is concerned, since they only differ by reason of the small variations between the respective modification factors we have been discussing. It is interesting to record that before van der Pol and Bremmer had published in their third paper the set of ground curves for λ from 1 m. to 20 km. for $\epsilon=4$ and $\sigma=10^{-13}$ e.m.u., and for $\epsilon=80$ and $\sigma=4 \times 10^{-11}$ e.m.u., we had independently computed a set of curves for λ from 2 m. to 2 km. for $\epsilon=5$ and $\sigma=10^{-13}$ e.m.u., and for $\epsilon=80$ and $\sigma=4 \times 10^{-11}$ e.m.u.

A comparison of these curves shows a close agreement over the whole range covered by our curves, both in slope and in absolute magnitude. We thus have a very satisfactory check on the practical equivalence of the methods. The small variations produced by the use of the different

approximations are probably within the limits of accuracy to which we are justified in working, when we remember the difficulty of deciding the value of the earth constants involved, and of determining the true radiated power in any particular case. Also as the wave-length decreases, the effect of uneven and inhomogeneous ground, and modifications produced by ionospheric reflexions and atmospheric refraction, will cause increasing divergences from the ideal case we are considering in our ground curves.

In particular we assumed that when $\sigma^{1/2}\lambda^{5/6} < 0.1 \times 10^{-7}$ (σ in e.m.u., λ in km.), the upper limit of β_s can be used, for which we took our value of 53.0; while for values of $\sigma^{1/2}\lambda^{5/6} > 0.1 \times 10^{-7}$ we used values of β_s computed on the assumption that $2\sigma\lambda c \gg \epsilon$. This bears out our statement that away from the upper limit $2\sigma\lambda c \gg \epsilon$ in any practical case, and this can be shown to hold even when $\sigma = 10^{-14}$ e.m.u. For very short waves the use of the upper limit for β_s of 53.0 instead of 55.0 only produces a small difference of slope, and in practice the field strength will have dropped well below a workable signal before any serious error has arisen with increasing distance from the transmitter, even allowing for subsequent height-gain effects.

7. The Height-Gain Factor.

Returning to the form of the height-gain factor in (2.9), we have given $\psi(r)$ in the form

$$\psi(r) = \frac{1}{\sqrt{\sin \alpha}} \cos (\Delta + q\pi),$$

where

$$\Delta = \frac{x\alpha^3}{3}$$

and

$$\alpha^2 = \frac{2h}{r_0} + \alpha_0^2.$$

h is the height of the transmitter above the ground, so that

$$h = r - r_0,$$

and α_0^2 is given by (5.2). Thus

$$\psi(r_0) = \frac{1}{\sqrt{\sin \alpha_0}} \cos (\Delta_0 + q\pi),$$

where

$$\Delta_0 = \frac{x\alpha_0^3}{3},$$

and

$$F = \frac{\psi(r)}{\psi(r_0)} = \sqrt{\frac{\sin \alpha_0}{\sin \alpha}} \cdot \frac{\cos (\Delta + q\pi)}{\cos (\Delta_0 + q\pi)} \dots \dots \dots (7.1)$$

Wwedensky derived his height-gain factor in exactly this form, except, of course, that he puts $\pi/4$ where we have used $q\pi$. He differs from us, however, in his method of handling this expression, as he uses a transformation in deriving the modulus of F_s which makes explicit use of the value $\pi/4$. On the other hand, we write

$$\frac{\cos (\Delta+q \pi)}{\cos (\Delta_0+q \pi)}=\cos (\Delta-\Delta_0)-\tan (\Delta_0+q \pi) \cdot \sin (\Delta-\Delta_0),$$

and irrespective of the precise value given to q , the proper value relation gives

$$\tan (\Delta_0+q \pi)=-\eta, \quad . \quad . \quad . \quad . \quad (7.2)$$

where η is given by (5.3). Thus we give

$$\frac{\cos (\Delta+q \pi)}{\cos (\Delta_0+q \pi)}=\cos (\Delta-\Delta_0)+\eta \sin (\Delta-\Delta_0).$$

Wwedensky has shown, by taking several particular examples, that under some conditions on short waves there is first of all a diminution of signal strength on leaving the surface of the earth. Our method of developing the height-gain factor verifies this conclusion, and gives a simple analytical expression for the position and magnitude of the drop. Our height-gain curves agree very well with the curves that Wwedensky has calculated, and we also find that the initial drop becomes negligible both on very short and on long waves. The height-gain factor is thus strictly equivalent on the two methods, but it appears that our transformation leads to a more flexible and powerful way of studying the detailed implications of the theory.

When we compare our height-gain factor with that given by van der Pol and Bremmer, it is a much more difficult task because of the Hankel form in which their factor is expressed. But it is in this comparison that the serious exception to the equivalence occurs to which we referred in the introduction. They express $\psi(r)$ as a ζ function, which may be written

$$\psi(r)=\zeta_{n_s}^1(kr),$$

where

$$k=\frac{2 \pi}{\lambda},$$

so that

$$F_s=\frac{\psi(r)}{\psi(r_0)}=\frac{\zeta_{n_s}^1(kr)}{\zeta_{n_s}^1(kr_0)},$$

where $kr_0=x$, and n_s is the proper value given by (2.2).

These ζ functions are then converted into Hankel functions of order $\frac{1}{3}$. The development of the functions for purposes of computation leads to

considerable difficulties, but they give two different types of expansion according as $h \leq 56\lambda^{2/3}$ or $h \geq 56\lambda^{2/3}$, where h and λ are in metres. This corresponds to being well below or well above the fictitious sphere of radius $r_0 + h_1$ discussed in our paper. It thus appears that, although our relation derived from the tangent approximation can be computed continuously from the ground up to great heights, the Hankel form of the height-gain factor is difficult to handle in the region of the height h_1 .

The expansion which they give for points near to the ground, i. e., for $h \leq 56\lambda^{2/3}$, does not appear to possess the generality of the form which our expression takes when h is very small, since, unlike ours, it does not involve the earth constants explicitly in terms of ζ . It is difficult to know exactly what limitation is implied in the condition $h \leq 56\lambda^{2/3}$, but if we approximate to their expression for F_s in the form $F_s = 1 + Bh$ for very small heights*, the value of B depends markedly on the higher terms of their expansion, which they only quote to three terms. It appears, therefore, that their factor for $h \leq 56\lambda^{2/3}$ is not very convenient for practical application and is not in a form in which direct comparison can be made with our factor for small values of h .

When we consider their expressions for $h \geq 56\lambda^{2/3}$, we find that they are simpler to handle, since they do not involve expansions in the form of a series. They find that the factor takes different forms according as $\sigma = \infty$, or $|\delta| \ll 1$. The latter case, corresponding to heavy absorption, implies that with practical values of σ , $\epsilon \gg 2\sigma\lambda c$; as we have seen in quoting the form of δ in (3.1), this means that ρ_s has its upper limit value, and that $\delta \propto \lambda^{1/3}$. Now the forms of their function F_s in these two cases are similar, remembering that ρ_s has its lower and upper limit value respectively, except for an extra factor, which can be written as T , in the case of $|\delta| \ll 1$. Using their τ_s instead of our ρ_s , T is given by

$$T = \frac{24}{5} \frac{(-2\tau_s)^{3/2}}{\left(1 - \frac{96}{5}\delta\tau_s^2\right)} \cdot \cdot \cdot \cdot \cdot \quad (7.3)$$

We shall see later that when $h > 10h_1$, our height-gain factor can be put in a simple approximate form. Comparison for the case $\sigma = \infty$ shows complete agreement with the factor given by van der Pol and Bremmer. For $|\delta| \ll 1$ we get a factor corresponding to T in (7.3), which from (7.1) is given by

$$T = \frac{[\cos(\Delta_0 + q\pi)]_{\sigma=\infty}}{[\cos(\Delta_0 + q\pi)]_{|\delta| \ll 1}} \cdot \cdot \cdot \cdot \cdot \quad (7.4)$$

* For comparison with our form $F_s = 1 + \left| \frac{2\pi e^{i(\frac{\pi}{2} - v)}}{\lambda} \frac{1}{|\zeta|} \right| \cdot h$

Now from (7.2) we have $\Delta_0 + q\pi = \tan^{-1}(-\eta)$. When $\sigma = \infty$, $\zeta = \infty$, and thus η from (5.3) is zero. Thus $|\cos(\Delta_0 + q\pi)|_{\sigma=\infty} = 1$.

When $|\delta| \ll |\eta|$, and $|\cos(\Delta_0 + q\pi)|_{|\delta| \ll |\eta|} = |1/\eta|$. Thus from (7.4)

$$|T| = |\eta|.$$

Putting in the value of η we get, by replacing $|\rho_s|$ by $|\tau_s|$,

$$|T| = \left| \frac{x^\dagger}{\zeta \sqrt{2\tau_s}} \right|.$$

By using (3.1) this may be written

$$|T| = \left| \frac{1}{\delta \sqrt{2\tau_s}} \right| \cdot \cdot \cdot \cdot \cdot \cdot \quad (7.5)$$

We see at once that this does not agree with the value of $|T|$ derived from (7.3). As $\lambda \rightarrow 0$, $\delta \rightarrow 0$, and our $|T| \rightarrow \infty$, whereas from (7.3) their

$$|T| \rightarrow \left| \frac{24}{5} (2\tau_s)^{3/2} \right|,$$

which is finite and of the order of 40 for $s=0$.

Now if in the factor $(1 - 96/5 \delta \tau_s^2)$ in the denominator of T in (7.3) the term 1 is omitted, then (7.3) leads to the same value for $|T|$ as our value in (7.5). Thus the whole question of the discrepancy turns upon the correctness or otherwise of including this term 1 in the denominator of T . It is not possible to see from their analysis how van der Pol and Bremmer have obtained this factor, since they only state the final formulæ to which the appropriate approximations to the Hankel functions lead. But it is difficult to know how by using the tangent approximation we should have introduced an error which gets progressively worse, in the region where van der Pol and Bremmer say that this approximation is very good, namely, when δ gets very small. There is nothing inconsistent with the idea that the height-gain factor approaches infinity as $|\delta| \rightarrow 0$, i. e., as $\lambda \rightarrow 0$, since at the same time the exponential attenuation factor $\exp\left(-\frac{\beta_s \phi}{\lambda^{1/3}}\right)$ approaches zero, and C in the denominator of A_s approaches infinity.

As it is a matter of considerable importance to resolve this discrepancy, we shall give two indirect arguments which suggest that van der Pol and Bremmer are in error, and that the term 1 should really be omitted. In the first place we notice that their height-gain factor contains as its denominator before expansion the function

$$H_{1/3}^1 \left\{ \frac{(-2\tau_s)^{3/2}}{3} \right\},$$

while in their later paper they show that the upper limit of τ_s as $\delta \rightarrow 0$ is given by the roots of the equation

$$H_{1/3}^1 \left\{ \frac{(-2\tau_s)^{3/2}}{3} \right\} = 0.$$

We should therefore expect that when δ is very small, but not zero, this Hankel function should be very small and have a value proportional to δ , and that the denominator in the height-gain factor F_s should contain a factor proportional to δ , and not involving it in a factor of the type $(1 - 96/5 \delta \tau_s^2)$.

Secondly, if we take the series of wave-lengths 7 m., 70 cm., 7 cm., and 7 mm., considered by van der Pol and Bremmer, and compute the height-gain factors for $\epsilon = 4$ and $\sigma = 10^{-13}$ e.m.u. as they have done, the discrepancy will only be slight on 7 m., but marked on 70 cm., and extreme on 7 cm.

and 7 mm. The curves which they give for $\left| \frac{\pi}{\pi_{pr}} \right|$ for $h_1 = 100$ m. and $h_2 = 0$, and for $h_1 = h_2 = 100$ m. on these wave-lengths should therefore be seriously modified according to us, the curves for 7 cm. and 7 mm. needing to be raised considerably. It is interesting to note that the curves which they give for these wave-lengths for $h_1 = 100$ m. and $h_2 = 0$ show pronounced kinks, and that they cross over and go below the curve for $\lambda = 0$.

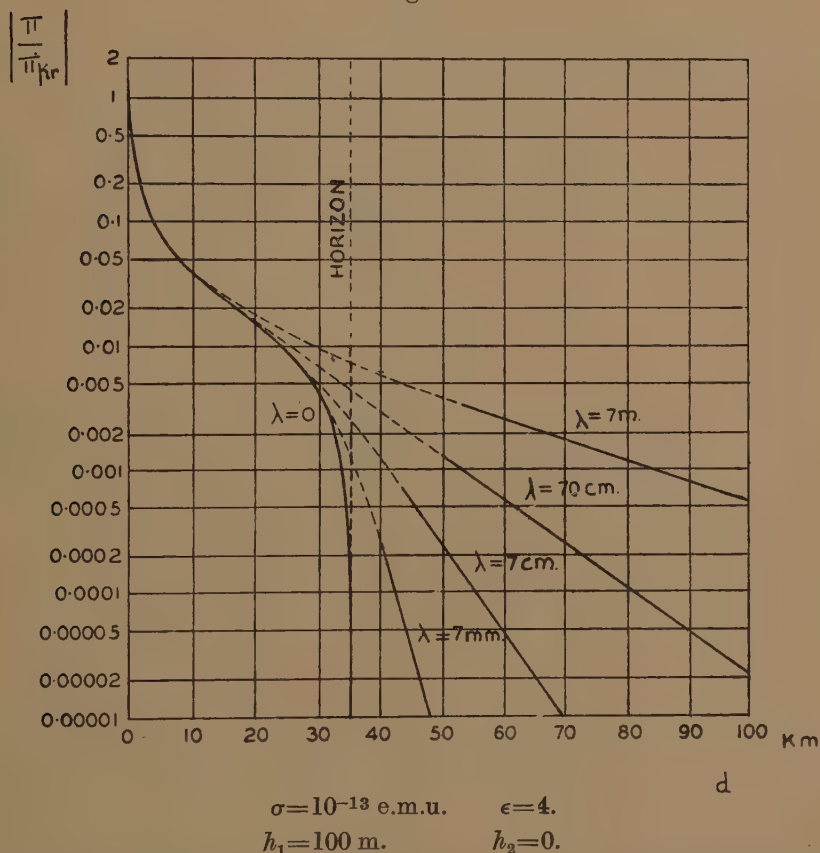
It is difficult to see the physical significance of these features, and when the curves are recalculated with our height-gain factor these anomalies disappear. This is seen in fig. 1, which should be compared with their curves to see the order of the difference between the height-gain factors for the very short waves. As we are only concerned with showing the essential difference between the two sets of curves, we have only used one term of the diffraction formula, and have interpolated by eye to join up with the curve for short distances computed by assuming a direct ray and a single reflected ray. Our curve for $\lambda = 0$, calculated on the simple reflexion picture, lies about 6 decibels above van der Pol and Bremmer's corresponding curve in the middle region between the initial and final rapid drops. This may be because we have assumed the flat earth reflexion coefficient instead of their spherical reflexion coefficient, though we should not have expected so big a difference on this account.

Our curves in fig. 1 seem more plausible than theirs, and show more convincingly the way in which the reduction of signal strength beyond the horizon becomes more rapid and approaches the optical case as the wave-length is decreased. We therefore feel that with the previous indirect argument this comparison suggests strongly that our form of

the height-gain factor is essentially correct, and that a definite error has arisen in van der Pol and Bremmer's analysis in this respect *.

While acknowledging the greater accuracy of the Hankel approximation in the general analysis, we consider that we are justified in using our modified tangent approximation for all practical applications. In addition to computing the set of ground curves to which we have referred above,

Fig. 1.



Curves for comparison with the curves given by van der Pol and Bremmer.

we have made many applications of our height-gain analysis covering a large range of heights and wave-lengths. In the course of this work we

* While this paper has been in the press the author has been in communication with Drs. van der Pol and Bremmer, and has been privileged to see a typescript copy of a further paper they have written in which this error is acknowledged and corrected.

have found various methods of manipulating the analysis for rapid computation.

We cannot go into this aspect of the problem in detail here, but as it will help to throw some additional light on the comparisons we have been making, we will proceed to consider the simple approximate forms which the height-gain analysis takes when $h \gg h_1$, and especially for short waves for which ρ_s can be assumed to have its upper limit.

8. Further Height-Gain Analysis for $h \gg h_1$.

In our previous paper the height-gain analysis was found to simplify considerably by the use of a parameter δ , which is a real positive angle. This δ must not be confused with the symbol δ used by van der Pol and Bremmer and quoted above in (3.1). On the ground δ has a value

$\delta_0 = \frac{2\pi}{3} + \omega$, and as h increases, δ decreases. The relation connecting δ and h is

$$h = h_1 \left[1 + \tan \left(\frac{\pi}{3} - \omega \right) \cot \delta \right], \quad \dots \dots \dots (8.1)$$

where

$$h_1 = r_0 \mid \rho_s \mid x^{-2/3} \cos \left(\frac{\pi}{3} - \omega \right), \quad \dots \dots \dots (8.2)$$

and is the height corresponding to the fictitious radius $\mid r_1 \mid$ discussed in our paper.

From (8.1) we see that when h reaches a value h_1 , δ has decreased to $\pi/2$, and $h > h_1$ corresponds to $\delta < \pi/2$. We now write (8.1) in the form

$$\tan \delta = \frac{h_1}{h'} \tan \left(\frac{\pi}{3} - \omega \right), \quad \dots \dots \dots (8.3)$$

where $h' = h - h_1$, and we see that δ is a function of h' , *i.e.*, in the diffraction theory we are essentially concerned with height referred to the fictitious sphere of radius $\mid r_1 \mid = r_0 + h_1$ as datum. This point was made in our discussion on the physical significance of the height-gain analysis, but it was not stressed in our presentation of the detailed form of the analysis.

We considered the general method of handling the height-gain analysis by means of the parameter δ , and in particular we showed that simple approximations can be used for points close to the ground for which $\delta_0 - \delta$ is a small angle. We will now consider the approximations which are obtained when h is much bigger than h_1 , so that δ is now a small angle. If we take as our condition that δ must be less than 10° , then (8.3) implies that $h > 10h_1$, when $\omega = 0$, which we will adopt as the criterion for the validity of our approximations.

Replacing $\tan \delta$ in (8.3) by δ , and h_1 by its value in (8.2), we have

$$\delta = r_0 \mid \rho_s \mid x^{-2/3} h'^{-1} \sin \left(\frac{\pi}{3} - \omega \right). \quad (8.4)$$

Returning now to our height-gain factor given in (7.1), we have shown that

$$\begin{aligned} \left| \sqrt{\frac{\sin \alpha_0}{\sin \alpha}} \right| &= \left[\frac{\sin \delta}{\sin \delta_0} \right]^{1/4} = \left[\frac{\delta}{\sin \left(\frac{\pi}{3} - \omega \right)} \right]^{1/4} \\ &= [r_0 \mid \rho_s \mid x^{-2/3} h'^{-1}]^{1/4}, \quad (8.5) \end{aligned}$$

by substituting for δ from (8.4).

The rest of the analysis simplifies somewhat if we take $\omega=0$ and give ρ_s its upper limit value, and as this is the case in which we are particularly interested we will make these assumptions. We therefore put $\omega=0$ in (2.7) and (8.4), and by using (2.3) we get

$$\delta = \frac{\beta_s}{2\pi} \cdot \frac{\lambda^{2/3}}{h'}, \quad (8.6)$$

where β_s has its upper limit value. Since β_s is by convention determined by expressing r_0 in km., λ and h' are here also in km., and for convenience we will express them throughout in km., as it is an easy matter to convert the formulæ into any other units of h and λ .

It is now convenient to express the height h' in terms of the horizon distance d' given by

$$d'^2 = 2r_0 h',$$

so that (8.6) may be written

$$\delta = \frac{\beta_s r_0 \lambda^{2/3}}{\pi d'^2}. \quad (8.7)$$

Now under our conditions $\eta \gg 1$, and we have shown that where $h \gg h_1$, we can then write

$$\begin{aligned} \left| \frac{\cos (\Delta + q\pi)}{\cos (\Delta_0 + q\pi)} \right| &= \frac{\mid \eta \mid}{2} \exp \left[\frac{1}{3} \left(2\rho_s \sin \frac{\pi}{3} \right)^{3/2} \operatorname{cosec}^{3/2} \delta \sin \frac{3}{2} \delta \right] \\ &= \frac{\mid \eta \mid}{2} \exp \left[\frac{1}{3} \left[\frac{2\beta_s}{(2\pi r_0)^{1/3}} \right]^{3/2} \frac{3}{2} \delta^{-1/2} \right]. \end{aligned}$$

Substituting for δ from (8.7) this becomes

$$\frac{\cos (\Delta + q\pi)}{\cos (\Delta_0 + q\pi)} = \frac{\mid \eta \mid}{2} \exp \left[\frac{\beta_s d'}{r_0 \lambda^{1/3}} \right]. \quad (8.8)$$

Thus for the total height-gain factor F_s in (7.1) we have, from (8.5) and (8.8),

$$|F_s| = \left[\frac{2r_0^2 \rho_s x^{-2/3}}{d'^2} \right]^{1/4} \frac{|\eta|}{2} \exp \left[\frac{\beta_s d'}{r_0 \lambda^{1/3}} \right].$$

Putting in the value of $|\eta|$ from (5.3) we get

$$\begin{aligned} |F_s| &= \frac{r_0^{1/2} x^{1/6}}{2^{5/4} |\zeta| d'^{1/2} \rho_s^{1/4}} \exp \left[\frac{\beta_s d'}{r_0 \lambda^{1/3}} \right] \\ &= \frac{(\sqrt{3})^{1/4}}{2^{3/2}} \frac{r_0^{1/2} (2\pi r_0)^{1/4} \lambda^{-1/6}}{|\zeta| d'^{1/2} \beta_s^{1/4}} \exp \left[\frac{\beta_s d'}{r_0 \lambda^{1/3}} \right] \\ &= \frac{457 \lambda^{-1/6}}{|\zeta| d'^{1/2} \beta_s^{1/4}} \exp \left[\frac{\beta_s d'}{r_0 \lambda^{1/3}} \right]. \quad \dots \quad (8.9) \end{aligned}$$

When both the transmitter and the receiver are raised we have heights h_T and h_R with corresponding horizon distances d_T and d_R , and for convenience we will regard these as being referred to the fictitious sphere, without using dashes to indicate the fact. The height-gain factor now becomes

$$|F_s| = \frac{209000 \lambda^{-1/3}}{|\zeta|^2 d_T^{1/2} d_R^{1/2} \beta_s^{1/2}} \exp \left[\frac{\beta_s}{r_0 \lambda^{1/3}} (d_T + d_R) \right]. \quad \dots \quad (8.10)$$

These factors can now be combined with the absolute value of the field strength on the ground, which from (2.4) and (5.7) is given by

$$E = \frac{0.047 |\zeta|^2 \lambda^{1/2}}{d^{1/2}} \exp \left[\frac{-\beta_s}{r_0 \lambda^{1/3}} d \right], \quad \dots \quad (8.11)$$

and for the field strength, when the transmitter or receiver alone is raised, we get from (8.9) and (8.11)

$$E = \frac{21.5 |\zeta| \lambda^{1/3}}{d'^{1/2} d^{1/2} \beta_s^{1/4}} \exp \left[\frac{-\beta_s}{r_0 \lambda^{1/3}} (d - d') \right]. \quad \dots \quad (8.12)$$

When both the transmitter and receiver are raised we get from (8.10)

$$E = \frac{9820 \lambda^{1/6}}{d_T^{1/2} d_R^{1/2} d^{1/2} \beta_s^{1/2}} \exp \left[\frac{-\beta_s}{r_0 \lambda^{1/3}} (d - d_T - d_R) \right]. \quad \dots \quad (8.13)$$

In deriving these expressions we have taken $r_0 = 6366$ km., and by our convention λ and d etc. are in km., while E is in mV/m. for 1 kW. radiated.

We are, of course, dealing here with the modulus of the individual terms of the diffraction formula. In the region where the first term does not predominate, and it is necessary to take several terms into account, we must know the relative phases of the various terms. It is not difficult to derive the complete phase value even in the general case, but in the

particular case which we are considering, when β_s has its upper limit value and δ is small, the expression simplifies. The final result, when all terms common to every term of the diffraction formula are omitted, may be put in the form Θ , where for the transmitter or receiver alone raised

$$\Theta = - \left\{ \frac{\sqrt{3}}{8} r_0 \rho_s x^{-2/3} + \frac{2\sqrt{2}}{3} r_0^{-3/2} x h'^{3/2} + \frac{\rho_s x^{1/3} d}{2r_0} + s\pi \right\}.$$

When both transmitter and receiver are raised we have

$$\Theta = - \left\{ \frac{\sqrt{3}}{8} r_0 \rho_s x^{-2/3} \left(\frac{1}{h_T} + \frac{1}{h_R} \right) + \frac{2\sqrt{2}}{3} r_0^{-3/2} x [h_T^{3/2} + h_R^{3/2}] + \frac{\rho_s x^{1/3} d}{2r_0} \right\} \quad \dots (8.14)$$

In these expressions we must remember that not only ρ_s , but also h' and h_T and h_R depend on the order s of the term, since $h' = h - h_1$ etc., and from (8.2) h_1 is a function of ρ_s . The term $s\pi$ arises from the height-gain factor, and is equivalent to the factor $(-1)^s$ given by van der Pol and Bremmer. In (8.14) this becomes $2s\pi$ and is therefore omitted for obvious reasons. The minus sign before these expressions for Θ can be omitted if we wish, and would not appear if we were to use the time factor $\exp(-i\omega t)$ instead of $\exp(i\omega t)$.

For our condition that $h > 10h_1$, it is therefore not a difficult process to evaluate the modulus of the vector sum of a number of terms of the diffraction formula, and to derive curves of the type given by van der Pol and Bremmer, in which the diffraction formula is computed for points well within the optical horizon. As the horizon is approached it is obvious that more and more terms of the formula must be included, and the number required for a given accuracy increases very rapidly within the horizon. We can see, however, in a general way from the form of E in (8.12) and (8.13) that, in spite of the increasing slowness of convergence, the series is still ultimately convergent.

The effect of the exponential part of the height-gain factor is to oppose the exponential attenuation factor in the ground curve formula, and it exactly counterbalances it on the horizon. But we see from (8.12) and (8.13) that this horizon is not referred to the sphere of the earth, but that each term of the diffraction formula is referred to its own fictitious sphere, since we are dealing with h' and not h . If for the moment we were to suppose that h' were replaced by h , then for all the terms the cancellation would occur at the same distance, corresponding to the optical horizon, irrespective of the value of β_s . Beyond the horizon, where the index becomes negative, the exponential terms would form a set of decreasing values due to the fact that β_s increases with the order s of the diffraction

term. Similarly, within the optical horizon the exponential terms would form an ascending series and the formula would become untenable.

Actually, however, for given values of the height above the ground, the corresponding h_T and h_R referred to the fictitious sphere gradually decrease with increasing s , since the height h_1 given by (8.2) is proportional to ρ_s , and therefore also increases with s . Thus by taking enough terms, even for points well within the optical horizon, we eventually reach a term for which the point is below the horizon referred to its own fictitious sphere, and the exponential terms then become a slowly descending series. Eventually, of course, we come to terms where the angle δ is no longer small, but goes through the value of $\pi/2$ at which h_1 has expanded to h , and still higher terms will approximate to their ground values. Where it is necessary to include some of the terms for which δ is no longer small, we must remember that our approximations are no longer valid, and that we must return to the complete form of the analysis.

We can see from the presence of the term $\lambda^{-1/3}$ in the exponential factor that, for a given accuracy, as λ decreases we can get nearer to the optical horizon by using one term only of the formula, but that within the horizon the number of terms required rapidly increases as λ decreases. Thus in drawing curves of the type in fig. 1, on the 7 m. curve we can compute the diffraction curve with sufficient accuracy with one term only for $d > 55$ km., whereas for the 7 mm. curve the single term is valid within 2 or 3 km. of the optical horizon at 35.7 km. As explained above we have not gone to the labour of computing the curves for more than one term, since we were only concerned in showing the order of the discrepancy between our height-gain factor and the one given by van der Pol and Bremmer. Their curves corresponding to our fig. 1 show how that within the optical horizon it is more and more difficult to compute the curves as the wave-length is decreased.

9. Comparison with Huyghens' Treatment.

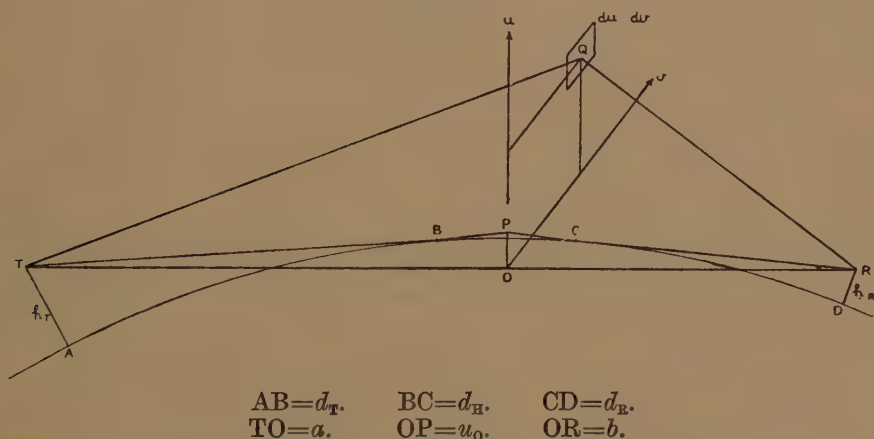
We notice that when the transmitter or receiver alone is raised the signal strength is still dependent upon the dielectric constant, since (8.12) contains $|\zeta|$, but that when both are raised it is independent of the nature of the ground, provided our condition $h > 10h_1$ is satisfied. In the latter case it appears that, due to the height of the transmitter and receiver, most of the energy reaching the receiver is diffracted down from the part of the wave front which clears the horizon, and that the intervening earth acts as a perfect absorber to the lower part of the wave. This is the simple picture that was used by Epstein⁽⁴⁾ in applying Huyghens' principle to the study of the diffraction of ultra short waves.

It is interesting, therefore, to compare the results of the complete diffraction analysis with those obtained by a simple application of

Huyghens' principle. Epstein was particularly interested in comparing his theory with the results of Marconi's experiments with a micro-wave beam, and he therefore took on his wave front a diffraction pattern corresponding to a plane wave emerging through a rectangular opening. In the case we are considering of a vertical dipole, we can treat the effective part of the wave front as if it were due to a point source, since we are dealing in practice with small angle diffraction.

In fig. 2, T is the transmitter at a height h_T above A, and AB is the corresponding horizon distance d_T . R is the receiver at a height h_R above D, and CD is its horizon distance d_R . TB and RC when produced cut in P, and, due to the effect of the intervening absorbing earth, only the part of the wave which is above P will produce a disturbance at R. We take a plane through P perpendicular to TR and cutting it at O, and define a set of rectangular axes Ou, Ov in this plane with Ou along OP.

Fig. 2.



Illustrating the application of Huyghens' principle.

The effect at R of an element $du.dv$ at Q, the point u, v , will be of the form

$$dE = k \cos \left(\omega t - \frac{2\pi}{\lambda} p \right) du.dv,$$

where k is a constant to be determined later, and

$$p = TQ + QR - TR.$$

If we write $TO = a$ and $OR = b$, then over the effective part of the wave front

$$p = (u^2 + v^2) \frac{(a+b)}{2ab}.$$

In the usual way we assume that the integral can be taken from $-\infty$ to $+\infty$ except that in the u direction the lower limit is u_0 , where $OP=u_0$. Thus at R we have

$$E = \int_{u_0}^{\infty} \int_{-\infty}^{\infty} k \cos \left[\omega t - \frac{\pi}{\lambda} \frac{(a+b)}{ab} (u^2 + v^2) \right] du \cdot dv.$$

We now put

$$u = \sqrt{\frac{\lambda ab}{2(a+b)}} \cdot u'_0, \quad \text{and} \quad v = \sqrt{\frac{\lambda ab}{2(a+b)}} \cdot v',$$

so that

$$E = \int_{u'_0}^{\infty} \int_{-\infty}^{\infty} \frac{k \lambda ab}{2(a+b)} \cos \left[\omega t - \frac{\pi}{2} (u'^2 + v'^2) \right] du' \cdot dv',$$

where

$$u'_0 = \sqrt{\frac{2(a+b)}{\lambda ab}} \cdot u_0. \quad \dots \dots \dots (9.1)$$

By using the Fresnel integrals this expression for E leads to an amplitude value of

$$E = \frac{k \lambda ab}{a+b} \cdot \frac{f(u'_0)}{\sqrt{2}}, \quad \dots \dots \dots (9.2)$$

where

$$f(u'_0) = \left\{ \left[\int_{u'_0}^{\infty} \cos \left(\frac{\pi}{2} u'^2 \right) du' \right]^2 + \left[\int_{u'_0}^{\infty} \sin \left(\frac{\pi}{2} u'^2 \right) du' \right]^2 \right\}^{1/2}. \quad (9.3)$$

We can now determine the value of k by expressing the fact that with the earth removed we should have the free-space value of $\frac{150}{a+b}$ mV/m. for 1 kW. radiated. We therefore put $u'_0 = -\infty$, whence $f(-\infty) = \sqrt{2}$, and from (9.2) we derive

$$k = \frac{150}{\lambda ab},$$

and substituting back in (9.2)

$$E = \frac{150}{a+b} \cdot \frac{f(u'_0)}{\sqrt{2}}. \quad \dots \dots \dots (9.4)$$

At the horizon when

$$u'_0 = 0, \quad f(u'_0) = \frac{1}{\sqrt{2}},$$

and

$$E = \frac{1}{2} \cdot \frac{150}{a+b},$$

i. e., the effect of half of the complete wave. Within the horizon we will get the usual rise through a series of diffraction maxima and

minima to the free-space value. Beyond the horizon the decrease in E is given by the reduction in the value of $f(u'_0)/\sqrt{2}$ as u'_0 is increased. In fig. 3 this function is shown up to $u'_0=3.75$, as computed from (9.3) with tables of the Fresnel integrals.

We now need the value of u'_0 . From the geometry of fig. 2, it is not difficult to show that

$$u_0 = OP = \frac{ab}{a+b} \cdot \frac{d_H}{r_0},$$

where $d_H = BC$, and represents the distance $d - d_T - d_R$ beyond the horizon. Thus from (9.1)

$$u'_0 = \sqrt{\frac{2ab}{\lambda(a+b)}} \cdot \frac{d_H}{r_0}.$$

Since we are only dealing with small angular distances we may write

$$a = d_T + \frac{d_H}{2}, \quad b = d_R + \frac{d_H}{2}, \quad a+b=d,$$

so that

$$u'_0 = \sqrt{\frac{2(d_T + d_H/2)(d_R + d_H/2)}{\lambda d}} \cdot \frac{d_H}{r_0} \dots \dots \dots (9.5)$$

$$= \frac{0.000222}{\lambda^{1/2}} \frac{d_H}{d} \sqrt{d_T d_R}, \dots \dots \dots (9.6)$$

when d_H is small in comparison with d_T and d_R .

Returning now to our expression in (8.13) we may write this for comparison with (9.4) in the form

$$E = \frac{150}{d} \cdot 8.82 \lambda^{1/6} \sqrt{\frac{d}{d_T d_R}} \cdot \exp \left[\frac{-0.00864}{\lambda^{1/3}} d_H \right],$$

in which we have given β_s the accurate Hankel value of 55.0 for $s=0$. By using (9.6) this may be written

$$E = \frac{150}{d} \cdot 8.82 K \exp [-39.0 K u'_0], \dots \dots \dots (9.7)$$

where

$$K = \lambda^{1/6} \sqrt{\frac{d}{d_T d_R}} * \dots \dots \dots (9.8)$$

If we write

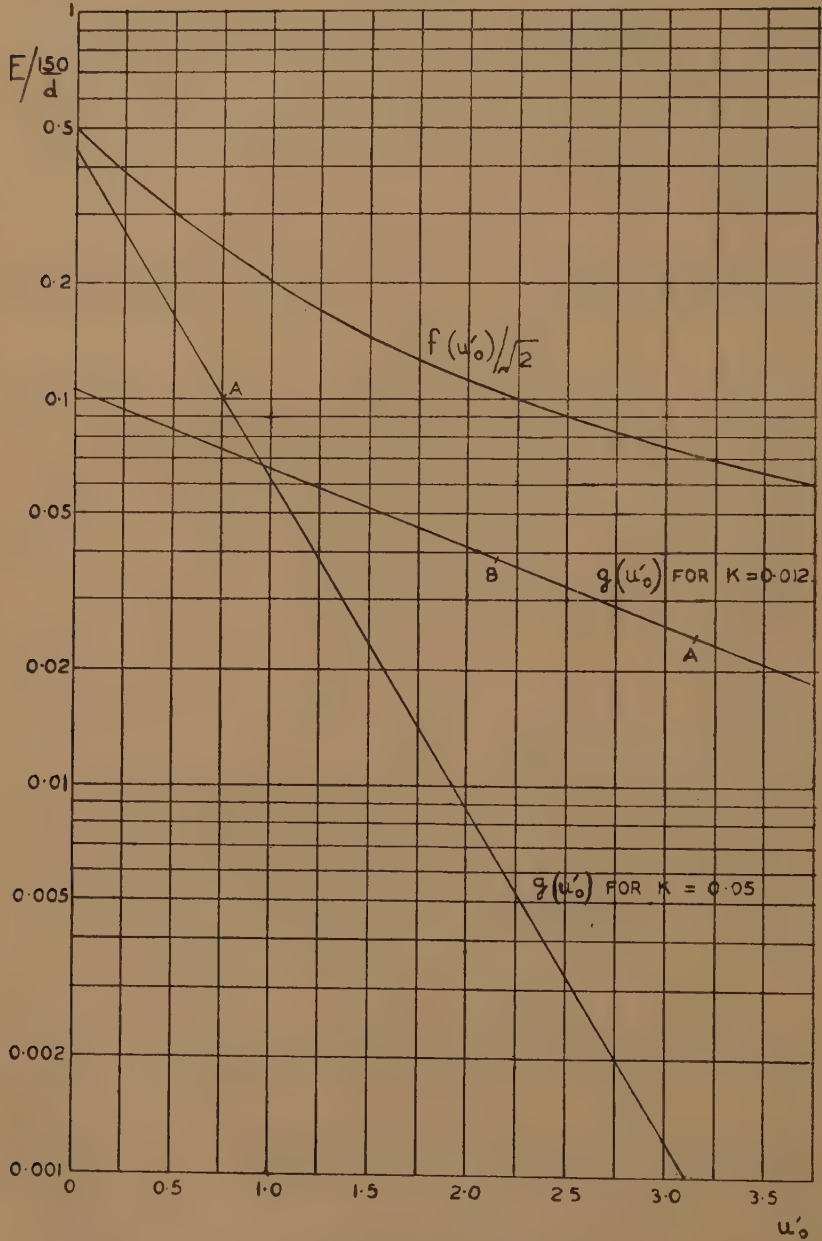
$$E = \frac{150}{d} \cdot g(u'_0),$$

where

$$g(u'_0) = 8.82 K \cdot \exp [-39.0 K u'_0],$$

* K can be considered constant since we are here using d_H as the variable and regarding it as small compared with the total distance $d = d_T + d_R + d_H$.

Fig. 3.



Curves of $E / \frac{150}{d}$.

$f(u'_0)/\sqrt{2}$ Huygens' principle.
 $g(u'_0)$ full diffraction theory (1st term).

for $g(u'_0)$ for $K=0.012$. The point corresponding to $z=1$ is shown as B, and since at $z=1$, $u'_0 = \frac{0.0256}{K}$, the point B always lies within the value $u'_0 = \frac{0.0378}{K}$ for the point A, beyond which by our convention the first term of the diffraction formula predominates.

All the cases satisfying the condition $h > 10h_1$ fall under the types of $g(u'_0)$ curve shown in fig. 3. At the point A, within which the first term no longer predominates, the curve obtained by the full diffraction analysis has approached towards the curve obtained on the simple Huyghens' theory. Within A the true curve for $g(u'_0)$, allowing for the higher terms, probably approaches the Huyghens' curve, and at the horizon would have a value of about 0.5. Within the horizon we should expect the system of interference maxima and minima given by the Huyghens' theory, and thus we may say that the theories agree in the neighbourhood of the horizon.

But as we go beyond the horizon the form of the curves becomes quite different. The Huyghens' curve becomes an inverse distance curve so long as d_H is small in comparison with d_T and d_R ; and then from (9.5) when d_H is large compared with d_T and d_R , so that d is approximately d_H , u'_0 becomes $\propto d^{3/2}$ and $f(u'_0)$ becomes $\propto d^{-3/2}$, giving from (9.4) $E \propto d^{-5/2}$. This is the result given by Epstein, and our Huyghens' analysis gives a result of exactly the same form as his. The full diffraction theory, however, gives a curve that is essentially exponential in character beyond the point at which the first term becomes predominant.

It thus appears that the simple Huyghens' treatment does not give the true form of the diffraction curve beyond the horizon, the true decrease of signal strength with increasing distance being much more rapid than it suggests. It is possible that the fundamental discrepancy is in some way connected with the simplifying assumptions made, and that a rigid application of Kirchhoff's rigorous expression of Huyghens' principle might reveal the exponential nature of the drop of signal strength with increasing distance.

Epstein uses the $d^{-5/2}$ law to explain the large signals obtained by Marconi with micro-waves beyond the horizon, and attributes the periods of weak reception to interference caused by atmospheric refraction. The full diffraction theory suggests that the opposite is actually the case, that the weak signals are of the order to be expected by diffraction alone, and that the effect of refraction is at times to aid reception considerably, while reducing the signals effectively to zero at other times.

In conclusion we may notice that the general equivalence which we have found in the neighbourhood of the horizon would not appear if the

height-gain factor of van der Pol and Bremmer were used. The equation corresponding to (8.13) for both transmitter and receiver raised would not be independent of $|\zeta|$, but would involve a factor $|\zeta|^2$, and as λ is decreased, the curve corresponding to our $g(u_0')$ would drop rapidly away from the value 0.5 in the neighbourhood of the horizon. This therefore provides another reason for doubting the correctness of their height-gain factor.

10. Appendix.

We take this opportunity of pointing out a feature of the phase integral analysis which was overlooked in our previous paper. Referring to section 2 of that paper we introduced the phase change on integration round the branch point as Ω , and showed by comparison with Watson's solution that Ω is approximately $\pi/2$. van der Pol and Bremmer have pointed out that in their analysis the tangent approximation gives this angle rigidly as $\pi/2$, without having to make any comparison with Watson's solution.

Actually, although we overlooked the fact, the phase integral theory also contains the value $\Omega=\pi/2$ in itself, as it is introduced by the use of the Jeffrey's second approximation which we adopted. We wrote this in the form

$$\psi_1(r) = \frac{1}{rf(r)} \exp(2\pi uS), \quad \dots \dots \dots (10.1)$$

where

$$\frac{\partial S}{\partial r} = \pm \frac{V}{\lambda} \quad \dots \dots \dots (10.2)$$

and

$$f(r) = V^{1/2}, \quad \dots \dots \dots (10.3)$$

in which

$$V = \left[1 - g(r) - \left(\frac{\lambda}{2\pi r} \right)^2 n(n+1) \right]^{1/2} \dots \dots \dots (10.4)$$

The \pm sign in (10.2) really belongs to the value of V in (10.4). We will then still have two values of S given by using the two values of V in

$$S = \frac{1}{\lambda} \int V dr,$$

which will be the same as we previously derived from (10.2). But we now get two values of $f(r)$ in (10.3) corresponding to

$$f(r) = \left\{ + \left[1 - g(r) - \left(\frac{\lambda}{2\pi r} \right)^2 n(n+1) \right]^{1/2} \right\}^{1/2}$$

and

$$f(r) = \left\{ - \left[1 - g(r) - \left(\frac{\lambda}{2\pi r} \right)^2 n(n+1) \right]^{1/2} \right\}^{1/2}.$$

Thus the two values differ by a factor $(-1)^{1/2}$, which is equivalent to a phase angle of $\pi/2$ between the two branches of $\psi_1(r)$ in (10.1). It is easy to show that with our conventions this corresponds to $\Omega = \pi/2$ on passing from the upgoing to the downcoming branch as defined by us. The phase integral involving the term $2\pi s + \Omega$ now contains the term $2\pi(s + 1/4)$ in agreement with van der Pol and Bremmer, and corresponding to $q = 1/4$.

11. References.

- (1) B. Wwedensky, Tech. Phys. U.S.S.R. ii. pp. 624-639 (1935) ; iii. pp. 915-925 (1936) ; iv. pp. 579-591 (1937).
- (2) B. van der Pol and H. Bremmer, Phil. Mag. (7) xxiv. pp. 141-176 (1937) ; (7) xxiv. pp. 825-864 (1937) ; (7), xxv. pp. 817-834 (1938). *Hochfrequenztechnik und Elektroakustik*, li. (6), pp. 181-188 (1938).
- (3) T. L. Eckersley and G. Millington, Phil. Trans. Roy. Soc. no. 778, cexxxvii. pp. 273-309 (1938).
- (4) P. S. Epstein, Proc. Nat. Aca. Sci. of U.S.A. xxi. pp. 62-68 (1935).

XLVIII. *A Kinematical Description of a Flat Space-Time.*

By C. GILBERT, Ph.D., King's College, Newcastle *.

[Received October 26, 1938.]

Summary.

The metric of the flat space-time is obtained by a simple transformation from the metric of special relativity. We find that there is a one-to-one correspondence between fundamental observers in these space-times, and that the transformation of coordinates is equivalent to a regraduation of an observer's clock-scale according to Milne's conventions. The geodesics in the space-time under discussion are found to correspond to certain of Milne's trajectories, which are supposed to give the paths of a statistical system of particles. These particles are shown to be acted on by an external force, derivable from a potential function which is a solution of Poisson's equation.

The Metric.

§ 1. WE write the metric of a space-time of special relativity in the form †

$$ds^2 = dt'^2 - dx'^2 - dy'^2 - dz'^2. \quad (1)$$

Transforming to polar coordinates

$$x' = r' \sin \theta' \cos \phi', \quad y' = r' \sin \theta' \sin \phi', \quad z' = r' \cos \theta', \quad . . . (2)$$

we get

$$ds^2 = dt'^2 - dr'^2 - r'^2(d\theta'^2 + \sin^2 \theta' d\phi'^2).$$

Now, if we make the transformation

$$t' = \tau(1 + \rho^2)^{\frac{1}{2}}, \quad r' = \tau\rho, \quad (3)$$

we get the metric

$$ds^2 = d\tau^2 - \tau^2 \left\{ \frac{d\rho^2}{1 + \rho^2} + \rho^2(d\theta'^2 + \sin^2 \theta' d\phi'^2) \right\}.$$

This describes the space-time Kermack and McCrea⁽¹⁾ considered when they carried out a comparison of Milne's system with general relativity.

On transforming the coordinates again according to

$$t = \frac{(1 + \rho^2)^{\frac{1}{2}}}{\tau}, \quad r = \frac{\rho}{\tau}, \quad \theta = \theta', \quad \phi = \phi', \quad (4)$$

* Communicated by Professor W. H. McCrea.

† It is assumed throughout the paper that the velocity of light is equal to unity.

we get the metric

$$ds^2 = \frac{1}{(t^2 - r^2)^2} \{ dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \} \dots \dots \dots (5)$$

Also, from (2), (3), and (4)

$$\left. \begin{aligned} t' &= \frac{t}{t^2 - r^2}, \quad x' = \frac{r \sin \theta \cos \phi}{t^2 - r^2} = \frac{x}{t^2 - r^2} \text{ (say),} \\ y' &= \frac{r \sin \theta \sin \phi}{t^2 - r^2} = \frac{y}{t^2 - r^2} \quad \text{,,} \\ z' &= \frac{r \cos \theta}{t^2 - r^2} = \frac{z}{t^2 - r^2} \quad \text{,,} \end{aligned} \right\} \dots \dots \dots (6)$$

where $r^2 = x^2 + y^2 + z^2$.

The metric (5) can then be written

$$ds^2 = \frac{1}{(t^2 - r^2)^2} (dt^2 - dx^2 - dy^2 - dz^2) \dots \dots \dots (7)$$

Since this metric has been obtained from (1) by means of the transformation (6), it will describe a flat space-time. We propose to study some properties of this space-time in the following work.

The Geodesics.

§ 2. We write the geodesic equations for the space-time of special relativity in the form

$$x' = u_0 t' + a_0, \quad y' = v_0 t' + b_0, \quad z' = w_0 t' + c_0 \dots \dots \dots (8)$$

Then the geodesics for the space-time described by (7) are found, from (6) and (8), to be given by

$$\left. \begin{aligned} x &= u_0 t + a_0(t^2 - r^2) \\ y &= v_0 t + b_0(t^2 - r^2) \\ z &= w_0 t + c_0(t^2 - r^2) \end{aligned} \right\} \dots \dots \dots (9)$$

In particular the paths of fundamental particles

$$\frac{x'}{u_0} = \frac{y'}{v_0} = \frac{z'}{w_0} \dots \dots \dots (10)$$

transform into the lines

$$\frac{x}{u_0} = \frac{y}{v_0} = \frac{z}{w_0} \dots \dots \dots (11)$$

The trajectories in the (x, y) plane correspond to trajectories in the (x', y') plane. Their equations are found by eliminating t between the first two of equations (9). If we write $A_0 = a_0 v_0 - b_0 u_0$, this gives

$$(A_0^2 - b_0^2)x^2 + 2a_0 b_0 xy + (A_0^2 - a_0^2)y^2 + A_0(v_0 x - u_0 y) = 0, \dots \dots (12)$$

provided a_0, b_0 are not both zero. This is the equation of a conic having real and distinct asymptotes if

$$(A_0^2 - b_0^2)(A_0^2 - a_0^2) < a_0^2 b_0^2, \\ \text{i. e., } A_0^2 < a_0^2 + b_0^2. \quad . \quad . \quad . \quad . \quad . \quad (13)$$

Now A_0 is the angular momentum of a particle in the (x', y') plane and therefore (13) is certainly true if

$$u_0^2 + v_0^2 < 1.$$

Also from (1), (7), and (8) we find that this condition is satisfied if

$$\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 < 1.$$

The trajectories of particles in the (x, y) plane having velocities less than the velocity of light are, therefore, *conics having real and distinct asymptotes*.

The angular momentum of a particle in the (x, y) plane is given by

$$A = x \frac{dy}{dt} - y \frac{dx}{dt};$$

using (9), we find that

$$A = -A_0 t^2 \frac{d}{dt} \left\{ t - \frac{x^2 + y^2}{t} \right\} = -A_0 t^2$$

when $x^2 + y^2 \ll t^2$.

The *angular momentum* of these particles therefore *varies as the square of the epoch*.

Kinematical Description.

§ 3. We have seen that the paths of fundamental particles in the space-time described by (1), transform into the straight lines (11) in the new coordinates. The number density of a system of fundamental particles in this space-time can be written⁽²⁾

$$n = \frac{Bt'}{(t'^2 - r'^2)^2},$$

where B is a constant. This number density satisfies the equation of continuity for these particles. Applying the transformation (6) we find that in the new coordinates

$$n = Bt(t^2 - r^2).$$

This is a solution of the equation of continuity in the space-time described by (7), for a system of particles following the paths (11).

The element of volume for the space-like surfaces is, from (7),

$$d\tau = \frac{dx dy dz}{(t^2 - r^2)^{3/2}}.$$

Hence the number distribution is

$$n d\tau = \frac{Bt}{(t^2 - r^2)^2} dx dy dz, \quad (14)$$

and there will be a one-to-one correspondence between the particles of this system and the fundamental particles in the space-time of special relativity.

§ 4. Our transformation from the space-time of special relativity is equivalent to a regraduation of the clock-scale of a fundamental observer. If t' and t denote clock readings of observers in space-times described by (1) and (7), we have

$$dt' = \frac{dt}{t^2}.$$

On integration this gives

$$t't' = -1. \quad (15)$$

In this relation we have neglected the constant of integration, which only shifts the zero of t' by a constant amount. Now, if we assume Milne's light-signalling conventions⁽³⁾, we find that the relation (15) leads to the transformation of coordinates

$$t' = -\frac{t}{t^2 - \mathbf{P}^2}, \quad \mathbf{P}' = \frac{\mathbf{P}}{t^2 - \mathbf{P}^2}, \quad (16)$$

where \mathbf{P}, \mathbf{P}' are the position vectors of an event at epochs t, t' respectively. This is equivalent to the transformation (6), for the minus sign is not material from the present point of view.

§ 5. We now suppose an observer to be attached to each particle of the system having the spatial distribution (14), and obtain a kinematical description of the geodesics (9) in the experience of one of these observers.

We shall use the following notation :

$$X = t^2 - \mathbf{P}^2, \quad Y = 1 - \mathbf{V}^2, \quad Z = t - \mathbf{P} \cdot \mathbf{V}, \quad \xi = \frac{(t - \mathbf{P} \cdot \mathbf{V})^2}{(t^2 - \mathbf{P}^2)(1 - \mathbf{V}^2)} \quad . (17)$$

where \mathbf{V} is the velocity of the particle at (\mathbf{P}, t) .

The equation of motion for a particle of a statistical system has been given by Milne⁽⁴⁾ as

$$\frac{d\mathbf{V}}{dt} = G(\xi)(\mathbf{P} - \mathbf{V}t) \frac{\mathbf{Y}}{X}. \quad (18)$$

The function $G(\xi)$ has a specified form depending on the space-velocity distribution of the system. It has also been shown by Walker⁽⁵⁾ that the geodesic equations of certain space-times of general relativity can be written in the form (18), if $G(\xi)$ be replaced by a function $G(X)$ which

The geodesic equations of the space-time (7) can be written in the form

$$\frac{1}{Y^{\frac{1}{2}}} \frac{d}{dt} \left(\frac{\mathbf{V}}{Y^{\frac{1}{2}}} \right) = - \frac{2}{X} \left(\mathbf{P} - \frac{\mathbf{VZ}}{Y} \right) \quad . \quad . \quad . \quad . \quad . \quad (29)$$

This can be deduced from equation (18) when $G(\xi) = -2$. Associated with (29) we have the time component equation

$$\frac{1}{Y^{\frac{1}{2}}} \frac{d}{dt} \left(\frac{1}{Y^{\frac{1}{2}}} \right) = - \frac{2}{X} \left(t - \frac{Z}{Y} \right) \quad . \quad . \quad . \quad . \quad . \quad (30)$$

Substituting from (29) and (30) in (27) and (28) we find that

$$\begin{aligned} \frac{\partial \chi}{\partial \mathbf{P}} &= \frac{m \xi^{\frac{1}{2}}}{X} \left(\mathbf{P} - \mathbf{V} \frac{Z}{Y} \right) + \frac{\mathbf{V}}{Y} \frac{d}{dt} (m \xi^{\frac{1}{2}}) = m \xi^{\frac{1}{2}} \left(\frac{\mathbf{P}}{X} - \frac{\mathbf{V}}{Z} \right), \\ - \frac{\partial \chi}{\partial t} &= \frac{m \xi^{\frac{1}{2}}}{X} \left(t - \frac{Z}{Y} \right) + \frac{1}{Y} \frac{d}{dt} (m \xi^{\frac{1}{2}}) = m \xi^{\frac{1}{2}} \left(\frac{t}{X} - \frac{1}{Z} \right). \end{aligned}$$

The potential can therefore be taken to have the value

$$\chi = m \xi^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

This type of potential has been considered by Milne⁽¹¹⁾ and has been shown to be a solution of Poisson's equation, and to lead to a value for the gravitational constant which is in agreement with observation.

Conclusion.

§ 9. The chief object of this paper has been to show that the geodesic equations of a particular space-time of general relativity can be put in the form of the integrals occurring in Milne's theory. Secondary considerations have been the properties of Milne's system which can be constructed in this space-time. This system provides a simple model of the universe in which a gravitational potential field occurs.

The transformation (16) is of interest because, in addition to its application in the present instance, it has been found to transform the most general universe considered by Milne into a universe of a similar type. The corresponding transformation of time-scales (15) is also rather of the type we might expect to exist between the atomic time-scales occurring in the cosmological theories of Milne and Dirac. The possibility of reconciling the two theories with each other by some such change in the clock-scales of the observers can, however, only be regarded as an interesting speculation.

The writer wishes to thank Professor W. H. McCrea for his assistance and guidance on many occasions.

References

- (1) Kermack and McCrea, *Monthly Notices R. A. S.* xciii. p. 519 (1933).
- (2) E. A. Milne, 'Relativity, Gravitation, and World Structure,' Oxford (1935).
Cf. p. 88. Much of the subsequent work is concerned with theory given in this book, which will be referred to as W.S. for convenience.
- (3) W.S. p. 29. Cf. also *Proc. Roy. Soc.* clviii. p. 344 (1937).
- (4) W.S. p. 177.
- (5) A. G. Walker, *Monthly Notices R. A. S.* xcv. p. 263 (1935).
- (6) W.S. p. 153, equations 51-53.
- (7) W.S. p. 180. The result follows from equation (21) when $G(\xi) = -2$.
- (8) W.S. p. 212.
- (9) W.S. p. 190.
- (10) Milne, *Proc. Roy. Soc.* clix. p. 531 (1937).
- (11) Milne, *Proc. Roy. Soc.* cliv. p. 42 (1936).

XLIX. *The Atomic Heat of Potassium.*

By L. G. CARPENTER and C. J. STEWARD,
University College, Southampton *.

[Received September 9, 1938.]

Introduction.

THE atomic heat of potassium has been measured by several workers from very low up to room-temperature, but at higher temperatures and in the liquid phase little is known.

Potassium is an interesting substance to investigate, as on theoretical grounds it is to be anticipated that in the alkali metals the electrons are degenerate in the Fermi-Dirac sense, and so they are unlikely to contribute more than a few per cent. to the total observed heat capacity. One is therefore probably dealing with a metal in which the atomic vibrations are responsible for practically the whole observed effect.

As the melting-point of potassium is only 63.4°C ., and we have found it possible to work from -70° to $+336^{\circ}\text{C}$., our measurements cover a fairly considerable temperature range in the liquid phase—a matter of some interest, since very little is known of the specific heat of liquid metals.

Description of Apparatus.

The calorimeter used was of the Nernst type as developed in this laboratory for high temperature work by Carpenter and Harle †. In this instrument the blackened platinum coil which plays the double rôle of heater and of thermometer is prevented from making direct contact with the calorimeter, but is linked to it thermally by radiation.

It was found that nickel was the only material suitable for a calorimeter to contain molten potassium, and that the calorimeter must be made entirely of nickel, as hard solder or brazing is attacked by potassium. Preliminary work showed that a piece of nickel foil of a total area about 100 cm^2 changed in weight by not more than half a milligram when kept in contact with molten potassium and its vapour for 40 hours at a temperature of about 300°C . Accordingly the calorimeter was a cylinder of nickel 7.5 cm. long and 5 cm. in diameter, having a conical

* Communicated by the Authors.

† Proc. Roy. Soc. A, cxxxvi. pp. 243–250 (1932), and Proc. Phys. Soc. xlv. pp. 383–399 (1932).

neck at the top, into which was ground an accurately fitting nickel plug. It was made by turning the conical neck and round plate which formed the top in one piece out of the solid and autogeneously welding this into an open-mouthed spun pot of the appropriate dimensions. This work was carried out with great care and skill by Messrs. Nick-El-Ectro of Birmingham. In the original design * of the high-temperature radiation calorimeter the coil was prevented from making electrical contact with the calorimeter by spacers of asbestos fabric; and as an additional precaution pieces of mica were interposed between coil and calorimeter walls. This tended to make the thermal linkage of coil and calorimeter by radiation less close than it might be. In the present calorimeter, therefore, the method of mounting the coil was modified so as to give maximum electrical insulation with a minimum interception of the radiation between blackened coil and blackened wall of calorimeter. To this end eight strips of mica were fixed at equal intervals radially to the sides of the calorimeter, the length of the strips being parallel to the length of the calorimeter. The outer edges of the strips were notched, and the platinum wire which was wound round the calorimeter lay in the notches in a strain-free and secure manner. The coil consisted of four bifilar turns of 42 S.W.G. wire, which had been blackened by bakelite varnish. Determinations of the α and δ of Callendar's formula from the resistances at the temperatures of melting ice, boiling water, and freezing zinc showed that up to 500° C. at any rate the effect of the carbonized varnish on the resistance of 42 S.W.G. platinum is negligible, as values of $\alpha = 38.98 \times 10^{-3}$ and $\delta = 1.446$ were obtained. It was thought that the thermal time constant of the coil might be reduced by rolling out the platinum into a ribbon, but it was found that, with a ribbon, so large a proportion of the thermal capacity of the coil arose from the carbonized varnish that ribboning did not appreciably reduce the time-constant.

Some idea of the thermal linkage between the coil and calorimeter may be gained from the fact that at temperatures of -150° , $+100^\circ$, and $+300^\circ$ C. the excess of temperature of coil above that of calorimeter, when a current of 5 milliamps was flowing through the coil, was about 0.25, 0.07, and 0.025° C. respectively.

The ground nickel plug in the conical neck of the calorimeter was pressed home by means of a nickel set-screw which was tapped through the horizontal member of a light nickel inverted U-shaped bridge, the vertical members of which were attached by small bolts to the sides of the calorimeter at two points on the opposite ends of a diameter. The method of filling the calorimeter *in vacuo*, putting in the ground plug and driving it home without allowing air to come in contact with the potassium, will be described in the next section.

* *Loc. cit.*

The calorimeter hung in a massive copper mantle, silver-plated on the inside to reduce radiation loss from the calorimeter. The design of the silver radiation shields, leads, and suspensions of the calorimeter, and of the massive copper mantle in which it was hung, was similar to that already* described. In the present case, however, the copper mantle was provided with a winding of 44 ohms of 40 S.W.G. platinum laid in grooves cut on its surface, and insulated therefrom with refractories. This platinum winding made it possible both to heat the mantle electrically when that was desirable in order to hasten the attainment of thermal equilibrium, and to observe any slight thermal drift of the mantle during an experiment. The mantle, with the calorimeter hanging inside it, was suspended within an evacuated drawn-steel tube, closed at the lower end by a welded-in steel plate and at the upper end by a waxed-on copper cap, through which the electrical leads were passed in pyrex bushes. The whole tube was evacuated, through a welded-in steel side tube, by means of an 02 Metropolitan-Vickers oil-diffusion pump, backed by a mechanical rotary pump. The degree of vacuum attained was measured by a lamp type G.E.C. Pirani gauge.

The lower part of the evacuated steel tube was surrounded by an electric furnace with a copper muffle, the temperature of which was controlled by a Cooke and Swallow thermostat †, in which the sensitive Weston relay was replaced by a galvanometer and photocell.

The resistance of the calorimeter coil was measured by a Callendar-Griffiths bridge, and that of the platinum coil in the mantle by means of a Post Office box of ratio 100 to 1, the third arm of the box being extended by means of a slide wire, so that it would measure down to 1/10 ohm. The electrical input to the calorimeter during an experiment was calculated from the readings of a carefully calibrated ammeter and voltmeter, and of a standard Synchronome clock. The duration of electrical input and the magnitude of the resultant temperature-rise varied in different experiments, but were usually of the order of 4 minutes and 1° respectively.

Filling of Calorimeter.

In filling the calorimeter the object was to distil very pure potassium directly into the calorimeter in such a way that it came into contact with nothing but nickel, and, having got the calorimeter full, to prevent contact between the potassium and any substance with which it reacts until the measurements of thermal capacity had been made.

As starting material Kahlbaum's purest potassium was used, and subjected to a preliminary distillation, the beginning and end fractions

* *Loc. cit.*

† J. Sci. Inst. vi. p. 287 (1929).

being rejected. At the end of this operation an amount of potassium about 20 per cent. more than was required to fill the calorimeter was obtained in a nickel pot. The pot and the calorimeter were sealed up inside a pyrex apparatus, so that the calorimeter could be filled by distillation *in vacuo*. The potassium which entered the calorimeter was prevented by nickel-sheet linings from coming in contact, in the liquid phase, with pyrex, so that the metal had no opportunity to pick up impurities by running down in drops over pyrex surfaces, with which, in fact, we found it had a slight reaction.

The distillation into the calorimeter took place in a good vacuum, the rise of potassium level in the calorimeter being followed by X-ray photographs taken at 10-minute intervals. The distillation was stopped at a level sufficient to allow for thermal expansion during subsequent high-temperature work, and when cool the apparatus was let down to nearly atmospheric pressure with purified argon. Carefully dried and purified medicinal paraffin oil was then sucked into the apparatus under the slight residual vacuum until the whole calorimeter was immersed and the pure potassium surface protected. At this point the calorimeter was cut free from the surrounding pyrex, removed and immersed in a bath of dried and purified petroleum ether (B.-P. range 100–120° C.), the outside washed clean, and the heavy paraffin in the space inside, above the potassium, replaced by the much more volatile petroleum ether.

The ground-nickel plug was now inserted into the calorimeter-neck, but prevented by a suspension of fine wires from slipping right home, and the calorimeter placed inside a cylindrical apparatus of steel called the "diver." The diver could be evacuated and had glass windows through which the nickel plug inside could be seen. It was also fitted with a screwdriver which could be manipulated from outside through pressure tubing. When the calorimeter had been placed inside, the diver was evacuated and the volatile petroleum ether evaporated off. The screwdriver was then used to tighten the set screw (in the horizontal member of the inverted U-bridge described in the previous section) down on to the nickel plug. The diver was then let down to atmospheric pressure with argon, and left until the space in the calorimeter above the potassium was filled with inert gas by leakage through the ground joint. Finally the diver was opened, the calorimeter was removed from the diver, and after the winding and silver shields had been replaced it was resuspended in the apparatus, ready for measurement.

The $C_p - C_v$ Correction.

Over the range from 0° C. to the melting-point the equation

$$C_p - C_v = \frac{\alpha^2 VT}{\kappa} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

has been used, where α , V , T , and κ are the cubic expansion coefficient, atomic volume, absolute temperature and compressibility respectively.

Below 0°C. , however, the equation

$$C_p - C_n = AC_p^2 T \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

has been used, because it is particularly applicable to low temperatures, and because in this region the variation of α and κ with temperature have not been measured, and, therefore, equation (1) is useless. The value of A has been chosen so as to make equations (1) and (2) give the same numerical value at 0° C. In the liquid phase we have been able to calculate $C_p - C_v$ from (2) at only one temperature (at which κ has been measured). As formula (2) cannot be expected to apply to the liquid phase, it has not been possible to compute the correction over an extended range of temperature. For the solid between 0° C. and the melting-point α has been taken * as 24.9×10^{-5} at 20° C., and its variation with temperature assumed to be linear and equal to 8 per cent. for a 50° rise, the temperature variation being derived from the mean of two results by Hagen †, and comparing well with the data of Bernini and Cantoni ‡, who give a variation of 7 per cent. over the same temperature range.

In calculating V, the density at 20° C. has been taken as 0.86 and the atomic weight as 39.096.

For κ Bridgman's value \parallel of 36.4×10^{-12} cm.²/dyne⁻¹ at 45° C. has been adopted. As Bridgman made measurements at low pressures at one temperature only, we have been obliged to use the results of Protz on the temperature variation of κ . He found ¶ an increase of 10 per cent. between the temperatures of 1° and 50.9° C., and we have assumed the correctness of this, though as regards absolute magnitude we have adopted Bridgman's value at 45° C.

There is only one measurement of κ in the liquid phase, viz., that of Protz, who obtained ** the value of 49.2×10^{-12} cm.²/dyne at 90.9° C. As it does not seem possible to make any reliable estimate of its temperature variation, $C_p - C_v$ has been computed from equation (1) at 90.9° C. only, using for α a mean coefficient †† of 29×10^{-5} over the range 62° to 150° C., and for V a density ‡‡ of 0.83.

* I.C.T. vol. i. p. 104.

† *Ann. der Phys. und Chem.* xix. 7, pp. 436–474 (1883).

† *Nuovo Cimento*, viii. p. 241 (1914).

§ I.C.T. vol. i. p. 104.

Landolt-Börnstein Tabellen, Erst Erg. p. 24.

Ibid.

** Landolt-Börnstein Tabellen, Erst Erg. p. 24.

†† I.C.T. vol. i. p. 102.

†† *Ibid.*

The values of $C_p - C_v$, calculated at 50° intervals for the solid and at 364° abs. for the liquid, are given in Table I.

Experimental Results.

The heat capacity of the empty calorimeter was determined from -75° to $+420^\circ$. Twenty-eight different determinations were made in this temperature range and a smooth curve was drawn through the points. Incidentally it was satisfactory to find that the specific heat of nickel, as deduced from this curve, when correction had been made for the silver shields, mica slips, etc., agreed as closely as could be expected (in view of the uncertain allowances to be made for the non-nickel parts of the calorimeter) with the results of Lapp* and of Grew†. The

TABLE I.

Absolute temperature.	$C_p - C_v$.
50	0.005
100	0.137
150	0.228
200	0.326
250	0.465
300	0.610
Solid at } melting-point	0.688
364	0.701

greatest divergence was about 4 per cent. and the average divergence considerably less.

The heat capacity of the full calorimeter was measured over the temperature range -70° to $+336^\circ\text{C}$., eighty-four individual determinations being made. From each of these individual values was subtracted the value for the empty calorimeter as read from the smooth curve at the same temperature. The resulting heat capacity was multiplied by the ratio of the atomic weight to the actual weight of potassium used, in order to obtain C_p , the atomic heat at constant pressure.

In Table II. the values of C_p so obtained are given. They have been plotted in figs. 1 and 2, together with the results of other ‡ observers who

* *Ann. Physique*, xii. p. 442 (1929).

† *Proc. Roy. Soc. A*, cxlv. p. 509 (1934).

‡ Landolt-Börnstein Tabellen.

TABLE II.

Absolute temperature.	C _p .	Absolute temperature.	C _p .
203.4	6.517	334.2	7.930
204.1	6.776	334.3	7.945
275.7	6.913	334.5	7.936
276.8	7.217	334.8	8.023
287.0	7.156	334.8	8.092
288.0	7.284	335.5	8.125
301.6	7.503	335.7	8.974
302.4	7.380	335.9	9.063
305.1	7.265	336.0	13.22
305.9	7.542	336.1	88.10
308.3	7.470	336.2	390.0
309.5	7.639	336.51	Melting-point.
310.7	7.533	336.9	8.038
311.5	7.586	337.3	7.979
311.6	7.689	337.7	8.30
312.2	7.599	337.8	8.058
312.5	7.614	338.1	8.016
313.3	7.631	338.2	7.934
317.2	7.711	338.2	8.101
318.3	7.685	340.1	8.012
322.8	7.804	354.7	8.090
323.7	7.777	355.8	8.053
324.5	7.627	376.7	8.027
327.7	7.744	377.8	7.932
328.7	7.815	396.6	7.839
328.9	7.785	397.4	7.955
329.5	7.958	417.8	7.871
329.9	7.914	419.0	7.958
330.4	7.904	436.4	7.900
330.6	8.014	437.3	7.804
330.8	8.090	455.1	7.938
330.8	7.993	455.9	7.828
330.9	8.016	476.7	7.804
331.2	8.012	477.4	7.815
331.5	7.778	500.6	7.854
331.5	7.787	501.6	7.923
331.8	7.884	525.1	7.670
332.0	7.796	526.0	7.698
332.4	7.952	575.0	7.774
332.7	7.978	575.8	7.809
333.1	7.901	575.9	7.753
333.5	8.053	605.9	8.019
		609.5	8.098

measured true atomic heats, as opposed to mean values taken over a large temperature range.

Fig. 1.

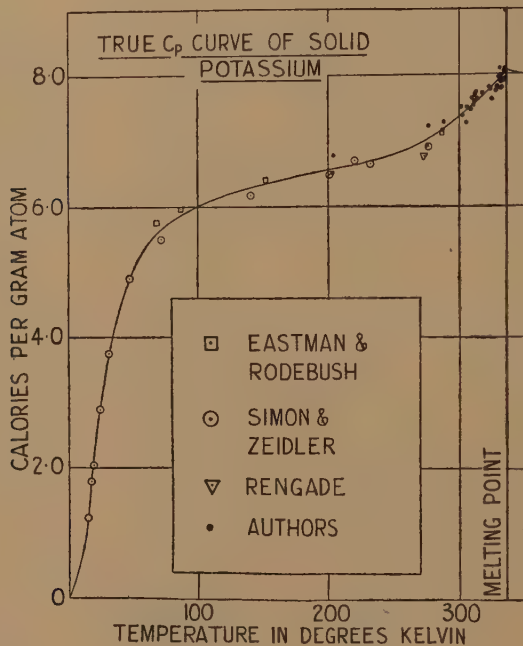
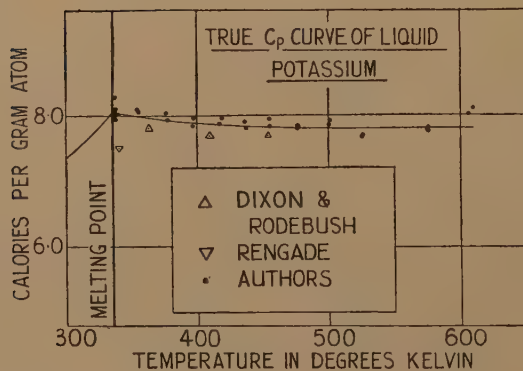


Fig. 2.



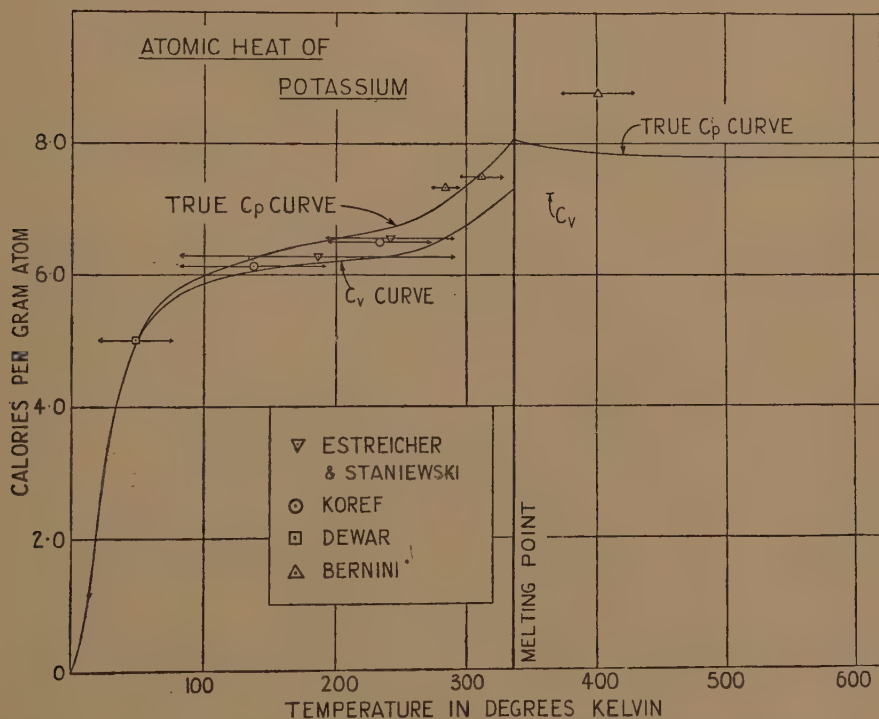
The last four points in the solid phase below the melting-point have not been plotted on fig. 1, because they cannot be conveniently represented on that scale. These high values are possibly due to pre-melting, caused

by a trace of impurity. Their significance is discussed in the concluding paragraph of the section on "The Freezing-point and Purity of Potassium."

On figs. 1 and 2 we have also drawn in the smooth curves which best represent, in our judgment, the results of all the measured values of true atomic heat, including our own.

On fig. 3 the smooth curves for true C_p have been repeated, and the corresponding C_v curves have been drawn, using the $C_p - C_v$ values given in Table I. Table III. gives a list of the values of C_p and C_v read

Fig. 3.



from the smooth curves, the values below 200 abs. being based entirely on the work of others. In addition, in fig. 3 the results have been plotted of observers* who determined mean atomic heats, the temperature ranges involved being shown by the length of the horizontal arrows.

Search for Transition Point.

It has been reported by Bidwell † that the electrical resistance and thermoelectric power against platinum show discontinuities at about

* *vide* Landolt-Börnstein Tabellen.

† Phys. Rev. xxiii. p. 367 (1924).

—120° C. To determine whether these changes are accompanied by thermal effects, a heating curve was taken through this region. With an argon atmosphere in the outer steel tube, the calorimeter and surroundings were cooled with liquid air. When the temperature of the calorimeter had become nearly stationary, the argon was pumped away, the liquid air replaced by a bath of solid CO₂ and methylated spirits, and the temperature of the mantle raised electrically to that of the bath, where it remained practically constant.

Readings of the platinum thermometer coil were then taken about every five minutes while the temperature rose from —130° to —105° C., the rate of rise being 0.2° per minute at the lower and 0.04°

TABLE III.

Solid potassium.			Liquid potassium.		
Absolute temperature.	C _p .	C _v .	Absolute temperature.	C _p .	C _v .
25	2.97	2.97	M.-P. 336.51	8.05	7.26
50	5.01	5.005	350	8.00	
75	5.71	5.62	364	7.96	
100	6.02	5.88	400	7.875	
150	6.33	6.10	450	7.82	
200	6.56	6.22	500	7.80	
250	6.79	6.32	550	7.80	
300	7.36	6.76	600	7.80	
M.-P. 336.51	8.06	7.37			

per minute at the higher temperature. The temperature-time graph so obtained was a smooth curve, and the greatest divergence of any point from the curve was not greater than 0.15°, while the average divergence was less than a quarter of this. Inspection of the curve failed to reveal any sudden change of slope, and the largest arrest which could possibly have occurred without detection corresponds to a latent heat of transition of about 1 calorie per gram atom. Hence, to the order of accuracy of our measurements no thermal evidence of Bidwell's transition point could be obtained.

Freezing-point and Purity of Potassium.

In order to determine the freezing-point of the sample of potassium used and to estimate its purity a cooling curve was obtained. To do this the calorimeter was heated to about 80° C. by means of its own

platinum coil and then allowed to cool, the outer steel tube being surrounded by ice. Under these conditions the mantle took up a steady temperature of about 7°C .

The coil resistance was read on the Callendar-Griffiths bridge about every five minutes as the freezing-point was approached, and much more frequently during supercooling (which was about 0.2°C .) and during the initial part of the freezing curve. The arrest was very sharp (apart from supercooling), and the total time of freezing was five hours. The curve, when examined in the way suggested by White*, gave a mid-point depression of 0.07°C ., and the maximum arrest temperature was about 63.35° . Hence from the cooling-curve measurements, assuming that the impurity does not form solid solutions, the freezing-point of pure potassium should be 63.42°C ., that of our sample being 0.07°C . lower, due to impurity. Assuming the freezing-point depressed according to Raoult's laws leads to a figure of 0.08 atomic per cent. for the impurity. In addition, a heating curve through the melting-point was obtained in order to gain further information as to the purity, the melting-point, and the specific heat of the solid and liquid phases in the immediate neighbourhood.

The mantle temperature was maintained about 5° above the melting temperature, and observations of the heating curve started when the calorimeter was 7° below the mantle. The variation of specific heat with temperature over the 2° immediately below the melting-point was deduced from the rate of temperature rise per unit difference of temperature between calorimeter and mantle, assuming that at the lower end of this range the specific heat was that measured, and recorded in Table II. Necessarily the variation of specific heat can be deduced in this way from a heating curve only quite roughly, but the experimental accuracy was sufficient to show that, starting from 2° below the melting-point, the specific heat rose at an increasingly rapid rate, being at 1° below the melting-point of the order of 25 per cent. greater than its initial value, and, at $\frac{1}{2}^{\circ}$, 160 per cent. It will be noticed that these values, though of the same order of magnitude, are higher than those recorded in Table II. The heating curve was continued until the whole of the potassium was molten. In order to hasten the melting process, measured amounts of heat were liberated electrically in the calorimeter by means of the platinum coil. Even so, the total time occupied by the melting process was about $6\frac{1}{2}$ hours.

Knowing the rate of heat supply by radiation, etc., from the mantle and the amount of heat supplied electrically, it was possible to calculate the fraction of potassium molten at various times during the $6\frac{1}{2}$ hours. The

* Jour. Phys. Chem. xxiv. p. 393 (1930).

temperatures of the potassium at these times were also known, and hence it was possible to construct a curve showing the equilibrium temperature of the solid-liquid system as a function of the fraction of the potassium melted. The difference between the temperatures at the middle and end of melting was 0.1°C ., which compares well with the value of 0.07° found for the mid-point depression of the cooling curve. The corresponding value of the melting-point of pure potassium was calculated as 63.40°C . The value from the cooling curve was 63.42° ; hence from the present work we deduce a mean value of 63.41° as the melting-point of pure potassium, assuming that any impurities present do not form solid solutions with the potassium.

The heating curve was continued for about 1° into the liquid phase in order to see if there was any abnormality in specific heat immediately above the melting-point.

The end of the melting curve was very sharp. The specific heat as deduced from the heating curve became constant within 0.05° of the temperature at which fusion was complete, and remained so within 10 per cent. (which was about the limit of accuracy of this measurement) until a degree above the melting-point, where we know (see Table II.) that it has attained the normal value for the liquid state.

It is very unlikely that the small amount of impurity present could sensibly affect the specific heat of the liquid phase; it remains, however, to consider the magnitude of the effect on the solid.

It is easy to deduce theoretically that any apparent rise in specific heat associated with increasing amounts of premelting (due to impurity) as the melting-point is approached should be proportional to the inverse square of the difference between the temperature in question and the true melting-point. This deduction is true even if the impurity can form both solid and liquid solutions, provided that the lines giving the composition of solid and liquid in the equilibrium diagram of the system potassium-impurity are straight. Reference to fig. 1 shows that the shape of the C_p curve above about 250°K . cannot be attributed to this cause, the rate of increase of C_p being much too small. On the other hand, the very rapid rise in C_p which occurs (see Table II.) within about a degree of the melting-point is most probably due to the small amount of impurity which the freezing and melting curves, already referred to, show to have been present. We therefore conclude the C_p of the solid phase was not sensibly influenced by impurity except within (say) a degree of the melting-point.

Determination of the Latent Heat of Fusion.

The latent heat of fusion was determined by supplying a measured quantity of energy to the full calorimeter, thereby raising it from 62.5°

to 67° C. in two hours. Correction was made for the amount of heat necessary to raise the potassium 4.5° (assuming the specific heat to be normal right up to the melting-point on both sides) and for the thermal capacity of the empty calorimeter over the same temperature range. It was also necessary to make a correction for the net heat exchange with the mantle during the experiment. This amounted to only about one per cent. of the total energy supplied.

The final result worked out to 0.568 kilo-calories per gram atom of potassium, with which may be compared the values * of 0.61, 0.53, and 0.50₄ obtained by Joannis, Bernini, Rengade, and Bridgman respectively.

Discussion of Results.

Figs. 1, 2, and 3 show that our determinations of C_p are in substantial agreement with those of other observers over that part of the temperature range which both we and they have covered.

The most significant part of our work, however, is concerned with the specific heat of the solid from room-temperature to the melting-point, where previously the only data had been two determinations by Bernini of the mean specific heat over a considerable temperature range.

It will be noticed in fig. 3 that from about 250° K. to the melting-point there is a large and increasing rise of the specific heat, both at constant pressure and at constant volume, and it has been shown in the section on "The Melting-point and Purity of Potassium" that this is very unlikely to be a spurious effect due to impurity.

Fig. 3 shows that in the region of 200° K. C_p has attained a magnitude slightly above the equipartition value of $3R$, and appears to be becoming constant, but that at 250° K. an anomalous rise sets in, continuing up to the melting-point.

It has been pointed out † that this phenomenon appears to be characteristic of the alkali metals. In the alkalis the electrons may be considered as completely degenerate in the Fermi-Dirac sense, and cannot contribute to the specific heat to any large extent. Hence the effect is probably to be ascribed to the lattice itself.

It may be that the crystal structure of the alkalis begins to break down under thermal agitation well before the melting-point itself, and this view is consistent with some work of Bidwell ‡ on the fading out of the X-ray diffraction patterns of lithium while still far from the melting-point. Investigation of solid diffusion in potassium as a function of temperature, using the radioactive isotope as an indicator, might throw some light on this.

* Landolt-Börnstein Tabellen.

† Carpenter, Harle and Steward, 'Nature,' cxli. p. 1015 (1938).

‡ Phys. Rev. xxvii. p. 381 (1926).

An alternative hypothesis (in principle distinct from that of lattice breakdown), which would explain the specific heat observations, is that of anharmonic oscillations, in which the mean potential energy exceeds the mean kinetic energy, on account of the restoring forces increasing less than linearly with the displacements at large amplitudes.

It does not appear possible on the present evidence to decide between these alternatives.

Our results in the liquid phase do not call for comment except to remark that when one subtracts the $C_p - C$ correction at 364°K. from the experimentally observed C_p , one obtains (see Table III.) a value of 7.26 calories per gram atom, which is definitely greater than $3R$, the value to be expected if in the liquid the atoms behaved as simple three-dimensional harmonic oscillators.

Acknowledgements.

We wish to express our thanks to the Government Grant Committee of the Royal Society and to the Research Committee of University College, Southampton, for grants towards the cost of apparatus used in this investigation.

L. Optical Dispersion and the Vibratory Doublet Photon.

By E. TAYLOR JONES, D.Sc., Professor of Natural
Philosophy in the University of Glasgow *.

[Received December 29, 1938.]

ONE of the most useful tests to which any theory of light can be subjected is that of its suitability for the explanation of the facts of optical dispersion in transparent media. In a theory of dispersion not only the nature of the incident light, but also the nature of its interaction with the particles of the medium, must be known or assumed with some precision in order that the velocity of propagation through the medium, and the manner of variation of this velocity with the frequency of the light, should be calculable. The classical theory of dispersion, in which the incident light is assumed to consist of continuous electromagnetic waves which preserve their frequency unchanged when they enter a dispersive medium, and cause the electrons attached to the atoms of the medium to make forced oscillations of this frequency, satisfies this test fairly well. The classical theory, however, fails to be convincing when, as in modern views, the light also is regarded as consisting of particles, *i. e.*, photons. For instance, there appears to be no more reason for supposing that the electrons make forced oscillations of the frequency of the photons than there is for assuming that the photons vibrate in the frequency of the electrons. Also the quasi-elastic forces which are assumed in the classical theory to exist in the atoms and to give the electrons in them natural frequencies of translational vibration, and the supposed frictional forces by which the natural vibrations are modified, have no place in more modern theories of atomic structure.

A theory of the photon recently put forward by the present writer † is based upon the properties of the electron regarded as a vibrator which can become coupled in a certain way with other electrons ‡. In this

* Communicated by the Author.

† Phil. Mag. xxiv. p. 458 (Sept. 1937).

‡ For shortness in reference the former papers on the vibratory electron will be referred to by numbers, as follows :—

- I. Phil. Mag. xxi. p. 337, Suppl. (Feb. 1936).
- II. Phil. Mag. xxii. p. 921, Suppl. (Nov. 1936).
- III. Phil. Mag. xxiv. p. 458 (Sept. 1937).
- IV. Phil. Mag. xxv. p. 682 (April 1938).

theory a photon is a doublet consisting of a positive and a negative electron coupled with each other so closely that the frequency of each is reduced to zero by the other. The doublet photon is therefore a particle of zero rest-mass, which, if free, has finite frequency and energy only when moving with the velocity of light *in vacuo*. It fulfils the conditions necessary for agreement with the Planck law of distribution of radiant energy, it has a transverse electrical axis as required for polarization, and it also has other fundamental properties identical with those possessed by light. In the present paper we shall enquire whether this theory of the photon also leads to a satisfactory dispersion formula for transparent substances.

As a preliminary to the discussion of dispersion, some further properties of vibratory doublets will first be described. In the former paper referred to * it was shown that the frequency ν of a free stationary doublet, which is given by the equation

$$n^2 = 4\pi^2\nu^2 \quad \dagger \\ = a - b, \quad \dots \dots \dots (1)$$

is zero because the two electrons forming the doublet are within the critical distance k/a of each other so that their electrostatic coupling b has the limiting value a . The coefficient k is the electrostatic coupling of two electrons at unit distance apart.

When the free doublet is moving with velocity v its pulsance becomes

$$n^2 = \frac{a - b}{1 - v^2/c^2} \ddagger, \quad \dots \dots \dots (2)$$

which, along with its mass, momentum and energy, is still zero unless $v = c$. In the latter case the expression for n^2 takes the indeterminate form $0/0$, the frequency of the doublet (now a photon) being determined by that of the source from which it was emitted §.

It might be thought that when the doublet is in translatory motion the two electrons composing it would be coupled magnetically as well as electrically. If this were the case the expression for n^2 would become

$$n^2 = \frac{a - b + bv^2/c^2}{1 - v^2/c^2} || \quad \dots \dots \dots (3)$$

* III. p. 459. See also I. p. 342.

† In this paper we shall use the term "pulsance" to denote $4\pi^2$ times the square of a frequency. Thus a is the pulsance of a free stationary electron, and $a - b$ is the pulsance of two coupled electrons of opposite kinds.

‡ III. p. 461.

§ III. p. 464.

|| IV. p. 685. It is here assumed that the law of the magnetic coupling, like that of the electrostatic coupling, is the same at very short as it is at moderate and great distances.

These differences of frequency would be accompanied by differences of wave-length and might be observed, if they are sufficiently large, as variations in the width of spectral lines depending upon the intensity of the light falling upon the slit of the spectroscope. Such effects, however, if they exist, must be very small, and we shall therefore neglect the coupling of photons in the same beam, and regard them as travelling independently of one another.

Turning now to the question of optical dispersion, we observe in the first place that when a photon is passing with velocity v near or through a system of stationary charged particles, *e. g.*, the electrons and atomic nuclei of a transparent medium through which the light is being transmitted, the photon becomes coupled with some of the fixed particles electrically, but not magnetically, and the common frequency of such coupled systems is to be formed from the group formula *

$$(n^2 - a_1)(n^2 - a_2) = \frac{b_1 b_2}{1 - v^2/c^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

or

$$n^2 = \frac{a_1 + a_2}{2} \pm \frac{1}{2} \sqrt{(a_1 - a_2)^2 + 4b_1 b_2 / (1 - v^2/c^2)}, \quad . \quad . \quad . \quad (10)$$

where a_1 is the pulsance of a typical particle of the fixed system before the arrival of the beam, a_2 that of a photon in the incident beam; b_1 is the electrostatic coupling and $b_1/(1 - v^2/c^2)$ the electrical coupling of the photon on a typical particle of the fixed system, and b_2 is the sum of the couplings of the fixed particles on the photon †. In equation (9) the quantities a_1 , a_2 and the product $b_1 b_2$ are essentially positive, but since the possibility that v^2/c^2 is greater than unity must be considered in any theory of dispersion, the quantity $1 - v^2/c^2$ may be positive or negative.

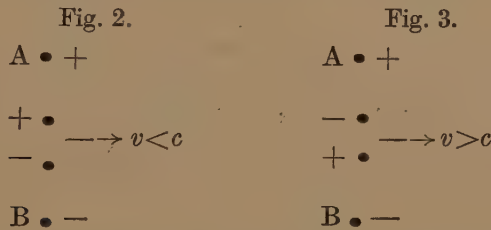
Equation (9) is applicable to any two electrically coupled systems of grouped electrons, but the number of types of solution of this equation, depending upon whether a_1 is greater or less than a_2 , and whether v^2/c^2 is less or greater than unity, is much restricted in consequence of the special properties of the photon which are represented by equation (2). In equation (2), which gives the frequency of a free photon travelling with the velocity of light, the numerator and denominator of the expression on the right-hand side are both zero. If, however, the photon

* II. p. 923; IV. p. 683.

† In strictness the group formula (9) requires that all the particles of each group should be similarly situated with respect to those of the other group. This is not likely to be the case in the present problem, but we may, without affecting the main argument, imagine the group of fixed particles to be replaced by a simpler group in which this condition is fulfilled.

passes near a fixed system of electrons which by their electrostatic action increase the electrostatic energy of the photon—the “field” of the fixed system in the position occupied by the photon must of course be non-uniform—the numerator in equation (2) becomes positive, and the denominator must also become positive, that is, v^2/c^2 must become less than unity, in order to keep the frequency real and finite. On the other hand, if the action of the field is such as to make the electrostatic energy of the photon negative, the denominator as well as the numerator in (2) must for the same reason become negative, that is, v^2/c^2 becomes greater than unity.

For the purpose of illustration the fixed system may be represented by the stationary electrons A and B in figs. 2 and 3, the photon passing mid-way between them. In accordance with the above restrictions, if the photon in passing through the fixed system has its axis placed as in fig. 2, its electrostatic energy is increased by A and B, and its velocity



must therefore be diminished. If, however, the photon passes as in fig. 3, its electrostatic energy is diminished and its velocity is increased.

Another restriction on the admissible types of solution of (9) seems reasonable when a_1 and a_2 are very different in value, *e. g.*, when light in the visual range is passing through a medium having an absorption band in the far ultra-violet or infra-red. It is improbable that a photon can produce a *great* change in the frequency of the much more massive fixed system which consists of a group of single electrons and nuclei *, and we shall accordingly exclude solutions of (9) in which n^2 differs greatly from a_1 , and admit only modes of vibration in which $n^2 - a_1$ is small in comparison with $n^2 - a_2$.

With these restrictions it is seen that such cases as

$$\left. \begin{aligned} a_1 \gg a_2, \quad v/c > 1, \\ n^2 < a_1 \gg a_2, \end{aligned} \right\} \quad \dots \quad (11)$$

* It may be remarked that the mass of a doublet travelling with velocity less than c is entirely due to the electrostatic influence of the charges near it. In the absence of such influence the mass of the doublet is zero.

the photon passing as in fig. 2, and

$$\left. \begin{array}{l} a_2 \gg a_1, \quad v/c > 1, \\ n^2 \gg a_1 < a_2, \end{array} \right\} \quad \dots \dots \dots (12)$$

the axis of the photon being as in fig. 3, though algebraically possible according to (9), must be excluded on physical grounds.

There are, in fact, only two cases left, if a_1 and a_2 are very different in value, which can be regarded as admissible, viz :—

$$\left. \begin{array}{l} \text{the higher frequency mode of the coupled} \\ \text{system of type fig. 2, in which } a_1 \gg a_2, \\ v/c < 1, \end{array} \right\} \quad \dots \quad (13)$$

and

$$\left. \begin{array}{l} \text{the lower frequency mode of fig. 3, in which} \\ a_2 \gg a_1, \quad v/c > 1. \end{array} \right\} \quad \dots \quad (14)$$

Taking the former of these two cases first, we find that since the right-hand side of (9) is positive, $n^2 - a_1$ and $n^2 - a_2$ must have the same sign. In the higher frequency mode n^2 is slightly greater than a_1 and much greater than a_2 . In these circumstances the rate of variation of n^2 with a_2 is also small. For, regarding a_1, b_1, b_2 in (9) as given, and $v^2/c^2 = f(a_2)$, we find

$$\frac{d(n^2)}{da_2} = \frac{n^2 - a_1}{2n^2 - a_1 - a_2} + \frac{b_1 b_2}{(1-f)^2 (2n^2 - a_1 - a_2)} \frac{df}{da_2} \quad \dots \quad (15)$$

$$= \frac{n^2 - a_1}{n^2 - a_2} + \frac{n^2 - a_1}{1-f} \frac{df}{da_2}, \quad \dots \dots \dots (15a)$$

if we neglect $n^2 - a_1$ in comparison with $n^2 - a_2$ in the denominator. If δa_2 is a given range of variation of a_2 small in comparison with $n^2 - a_2$, and $\delta n^2, \delta f$ are the accompanying variations of n^2 and f , we have from (15a)

$$\frac{\delta n^2}{n^2 - a_1} = \frac{\delta a_2}{n^2 - a_2} + \frac{\delta f}{1-f} \quad \dots \dots \dots (16)$$

The first term on the right of (16) is a small fraction, and in the second term $1-f$ cannot approximate to zero unless $b_1 b_2$ is evanescent. As to δf , since the whole of the possible range of variation of f in the mode (13) is only from 0 to 1, its variation over the limited range δa_2 of values of a_2 must be a small fraction of unity. Consequently, over such a range of frequencies the variation of $n^2 - a_1$ is small in comparison with $n^2 - a_1$, which may therefore be regarded as approximately constant.

Denoting $b_1 b_2 / (n^2 - a_1)$ by the single symbol D , we have, therefore, from (9)

$$1 - \frac{v^2}{c^2} = \frac{D}{n^2 - a_2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

or approximately,

$$1 - \frac{v^2}{c^2} = \frac{D}{a_1 - a_2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

We have arrived at equation (18) on the supposition that only one photon is passing through a fixed group of particles of the medium. If two or more photons, all having the same initial pulsatace a_2 , become coupled with the group simultaneously in a similar way, each of them will have its velocity diminished in accordance with equation (2), and each will have its effect on b_1 and $n^2 - a_1$, while b_2 remains unaltered and $n^2 - a_2$ is still approximately equal to $a_1 - a_2$. Equation (18) therefore still holds with the value of D unchanged, *i. e.*, D is independent of the intensity of the light. The ratio c/v is the refractive index of the group, and denoting this by μ , we have

$$1 - \frac{1}{\mu^2} = \frac{D}{a_1 - a_2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

from which

$$\mu^2 - 1 = \frac{D}{a_1 - a_2 - D}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

With the exception of the constant term in the denominator, the expression (20) for $\mu^2 - 1$ is of the same form as the dispersion formulæ of Sellmeier and Drude for a medium having a single absorption band in the ultra-violet. Expressed in terms of wave-lengths, λ being the wave-length of the incident light, λ_v that of light *in vacuo* of pulsatace a_1 , the equation becomes

$$\mu^2 - 1 = \frac{D \lambda_v^2 \lambda^2}{4\pi^2 c^2 (\lambda^2 - \lambda_v^2) - D \lambda_v^2 \lambda^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

the second term in the denominator $D \lambda_v^2 \lambda^2$ being a positive quantity smaller than the first.

In the other of the two admissible modes, *i. e.* (14), in which $a_2 \gg a_1$, $v/c > 1$, the right-hand side of (9) is negative, and $n^2 - a_1$ and $n^2 - a_2$ have opposite signs. The value of n^2 is therefore between those of a_1 and a_2 . In the higher frequency mode n^2 is slightly smaller than a_2 , and in the lower frequency mode n^2 is a little greater than a_1 , the latter mode alone being admissible. From equation (9) we have

$$1 - \frac{1}{\mu^2} = \frac{b_1 b_2}{(n^2 - a_1)(n^2 - a_2)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

in which $n^2 - a_1$ is again approximately constant over a range of values of a_2 small in comparison with $a_2 - a_1$. Denoting $b_1 b_2 / (n^2 - a_1)$ by D' , and taking $a_1 - a_2$ as approximately equal to $n^2 - a_2$, we have, therefore,

$$1 - \frac{1}{\mu^2} = \frac{D'}{a_1 - a_2}, \quad \dots \dots \dots (23)$$

and hence

$$\mu^2 - 1 = \frac{D'}{a_1 - a_2 - D'}, \quad \dots \dots \dots (24)$$

which is of the same form as (20).

When a photon is being transmitted through a transparent substance it will come under the influence of a large number of atomic groups in which, though the value of a_1 may be the same for all, the coupling product $b_1 b_2$, and consequently the value of D or D' may vary considerably from one to another. In equation (20) D is to be regarded as the average value of this coefficient for all the groups through which the photon passes with diminution of its velocity *. Similarly D' in (24) is the average value for those groups in which the photon comes under the influence of the low frequency vibration of the medium. The lower pulsantance of the medium being now denoted by a_1' , and the photon being supposed to come under the influence of both frequencies, the refractive index of the medium for light of free pulsantance a_2 is given by an equation of the form

$$\mu^2 - 1 = \frac{D_v}{a_1 - a_2 - D_v} + \frac{D_r}{a_1' - a_2 - D_r} \dots \dots \dots (25)$$

In terms of wave-lengths (25) assumes the form

$$\mu^2 - 1 = \frac{D_v \lambda_v^2 \lambda^2}{4\pi^2 c^2 (\lambda^2 - \lambda_v^2) - D_v \lambda_v^2 \lambda^2} + \frac{D_r \lambda_r^2 \lambda^2}{4\pi^2 c^2 (\lambda^2 - \lambda_r^2) - D_r \lambda_r^2 \lambda^2}, \quad \dots \dots (26)$$

λ_r being the wave-length *in vacuo* of light of pulsantance a_1' .

Equation (26) may also be written

$$\mu^2 - 1 = E \left(1 - \beta \frac{\lambda_v^2}{\lambda^2} \right)^{-1} - F \lambda^2 \left(1 - \gamma \frac{\lambda^2}{\lambda_r^2} \right)^{-1}, \quad \dots \dots (27)$$

where

$$E = \frac{D_v \lambda_v^2}{4\pi^2 c^2 - D_v \lambda_v^2},$$

$$\beta = \frac{4\pi^2 c^2}{4\pi^2 c^2 - D_v \lambda_v^2}.$$

* More exactly, μ^2 is the reciprocal of the average value of v^2/c^2 for all the groups with which the photon becomes coupled, and for the spaces between them where $b_1 b_2 = 0$ and $v = c$.

The quantity $D_v \lambda_p^2$ is smaller than, but comparable with, $4\pi^2 c^2$, so that E and β are not very different from unity. Similarly,

$$F = \frac{D_r}{4\pi^2 c^2},$$

$$\gamma = \frac{4\pi^2 c^2 - D_r \lambda_r^2}{4\pi^2 c^2},$$

so that F is of the order $1/\lambda_r^2$ and γ is rather less than unity.

When expanded, the two terms on the right of (27) give a descending power series in λ^2 beginning with a negative term in λ^2 . Such a series has been compared with experimental values of μ^2 for quartz, rock-salt, water, and other transparent substances *. Usually four terms of the series are found to be sufficient to give a good representation of μ^2 over a range covering the visible spectrum, the four coefficients involved being determined from observed values of μ^2 for four known wave-lengths.

It may be noticed that the thermal or other non-electrical energy of the medium is drawn upon when a photon enters it in circumstances similar to those of fig. 2, for the electrostatic energy of both the medium and the photon is increased by the coupling between them. In transmission of the type represented in fig. 3 the medium experiences an increase, the photon a diminution, of electrostatic energy, but these energy changes are not in general equal in magnitude, so that again the non-electrical energy of the medium becomes involved. Such changes are, however, only temporary. The principles of energy and momentum must hold throughout the process of transmission, and both the medium and the photon must be restored to their original state as regards energy when transmission is complete. No thermal effect is therefore to be expected in a material medium through which light is passing unless absorption occurs.

A necessary preliminary to absorption is that the photon should be brought to rest in one of the atomic groups of the medium, but the process of absorption is presumably not complete until the photon becomes free and is moving in the intermolecular spaces of the substance with a velocity less than c , having given up the whole of its energy to the medium. The condition that absorption should be possible is therefore $v/c=0$ †, and this is clearly possible only when the coupling is as in fig. 2, n^2-a_1 and n^2-a_2 being both positive quantities. From (9) therefore we have as the condition for absorption

$$(n^2-a_1)(n^2-a_2)-b_1b_2=0. \quad . \quad . \quad . \quad . \quad . \quad (28)$$

* See Wood, 'Physical Optics'; Houstoun, 'A Treatise on Light'; Gehrccke, 'Handbuch der physikalischen Optik.'

† The same condition, if it occurs at the surface of the medium, is evidently necessary for reflexion.

This condition holds for all values of a_1 and a_2 provided a_2 is not greater than a_1 , but the probability of this condition being satisfied depends upon the value of the coupling product b_1b_2 . In a former paper * it was shown that ratio of the amplitudes of vibration of two coupled systems (here the photon and the fixed group) is

$$\frac{B}{A} = \frac{b_2}{n^2 - a_2}, \quad (29)$$

and that when this ratio differs from both of the values ± 1 the two systems cease to oscillate continually in their common mode and at intervals make "partial oscillations" in their own free modes, the duration of the free oscillations diminishing as B/A approaches one of the above values. In the system of fig. 2, if a_2 is much smaller than a_1 , and, therefore than n^2 , and if b_2 is also small, B/A is small and positive, in which case a photon on its way through the transparent substance will encounter many of the atomic groups without becoming coupled with them, and will pass through them without change of its frequency or velocity. If in some of the groups the photon experiences larger values of b_2 , and therefore also of b_1 , the ratio B/A is greater in these groups, and the chance of the photon becoming coupled with them is increased. Also it appears from equation (2) that a large diminution of v^2/c^2 requires in general a correspondingly great increase of the numerator, that is, a larger value of b_2 .

It may be concluded that the chance of a photon being brought to rest by one of the atomic groups is greater the greater the value of b_1b_2 . Evidently appreciable absorption can occur with a_2 much smaller than a_1 if the coupling between the photon and some of the groups is sufficiently close. This is the general absorption always observed when light in the visual range is transmitted through a transparent substance, and according to this theory it is due to the influence of natural vibrations of the medium in the ultra-violet region.

A further increase of B/A , and therefore also of the chance of absorption, occurs when the denominator of (29) is diminished by increase of a_2 , *i. e.*, by raising the frequency of the incident light towards equality with that of the medium. For given values of b_1 and b_2 the ratio B/A reaches its greatest value when $a_2 = a_1$, giving rise to selective absorption which, on this view, is due, not to resonance, but to the greatly increased probability of a photon being brought to rest within the medium. This probability has its greatest possible value for a given medium if also $b_1 = b_2$, the ratio B/A being then unity.

* II. pp. 923, 936.

The general expression for μ^2-1 , for photons transmitted through a medium having one absorption band, is represented by the average of a number of terms similar to that on the right of the equation

$$\mu^2-1 = \frac{b_1 b_2}{(n^2-a_1)(n^2-a_2)-b_1 b_2}, \quad \dots \dots \dots (30)$$

which shows the effect of one atomic group of the medium. The principal feature of the variation of μ^2 with a_2 , or of refractive index with frequency of the incident light, at the absorption band, is sufficiently indicated by equation (30). As a_2 is increased through the value a_1 the product $(n^2-a_1)(n^2-a_2)$ suddenly changes from positive to negative, and μ^2 therefore changes from a value greater to one less than unity. Negative values of μ^2 do not occur, since the right-hand side of (30) cannot be less than -1 . The imaginary values of μ which appear in the classical theory do not occur in the present one, and the frictional forces of the former theory also are not required. The present theory, however, differs fundamentally from the classical theory in its view regarding the nature of light and its interaction with the particles of the medium.

The theory of the photon adopted in the present paper follows naturally from the idea that two mutually attracting or repelling electrons are a pair of coupled vibrators, the only assumption involved being that of a certain law of the coupling. No other particles are assumed or required to take any part in the theory. In former papers it has been shown that the fundamental facts of electrostatics and magnetism can also be satisfactorily accounted for in terms of the vibrating electron.

Natural Philosophy Department,
The University, Glasgow.
December 1938.

LI. *On some Geophysical Consolidation Problems.*

By H. Löwy, Dr.Phil.*

[Received October 1, 1938.]

ACCORDING to the theory of Terzaghi †, the settlement ("Setzung") s of a clay layer of the initial thickness h is equal to the height of the column of pore water, pressed out per unity of surface. Designating by v the water volume, contained in 1 cubic cm. of the clay, we put :

$$s = h \cdot \Delta v. \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The settlement can be measured by levelling ‡.

The clay, containing water in its pores, constitutes, from the electrical point of view, a metallic suspension, that is a mixture of electrically conductive particles (water) dispersed in an insulating substance (clay). One can, therefore, apply the general equations § which I have indicated for such suspensions. Designating by ϵ the dielectric constant of the suspension (the mixture), by ϵ_c the dielectric constant of the clay-substance, and putting, for abbreviation,

$$\frac{\epsilon - 1}{\epsilon + 2} = [\epsilon]$$

we have, according to these equations, in our special case ||,

$$[\bar{\epsilon}] = v + (1 - v)[\epsilon_c] \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and

$$\Delta[\bar{\epsilon}] = \Delta v(1 - [\epsilon_c]), \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

ϵ_c being independent of v .

Substituting the value of Δv from (3) to (1), we obtain the formula for the electrodynamic determination of settlement :

$$s = h \cdot \frac{\Delta[\bar{\epsilon}]}{1 - [\epsilon_c]} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In this equation h and ϵ_c represent known quantities, independent of the time.

* Communicated by the Author.

† K. v. Terzaghi and O. K. Fröhlich 'Theorie der Setzung von Tonschichten' (Leipzig u. Wien, 1936).

‡ K. v. Terzaghi, *Bautechnik*, xi. p. 579 (1933).

§ H. Löwy, *Phil. Mag.* xxvi. p. 453 (1938).

|| H. Löwy, *Phys. Zeits.* xxxv. p. 745 (1934).

In Table I., in the columns 1 and 2, are given, according to Terzaghi *, the variations of the pore-number p for different sorts of clay ("Porenziffer")

$$p = \frac{v}{1-v},$$

corresponding to a variation of pressure from 0 to 3 kg./cm.² Column 3 contains the variations of v , column 4 the values s/h , calculated according to (1). The corresponding variations of $\bar{\epsilon}$, calculated according to formula (2) for $\epsilon_c=2$ are given in the last column.

TABLE I.

Substance.	p .	v .	s/h .	$\bar{\epsilon}$.
Mississippi clay (New Orleans)	2.05-1.18	0.67-0.54	0.13	10-6.6
Clay, Columbia, S.A.	1.56-0.88	0.61-0.47	0.14	8.3-5.5
Mississippi, Gumbo, 50 m. depth	1.40-0.79	0.59-0.44	0.15	7.7-5.2
Mississippi Gumbo, 85 m. depth	1.42-0.72	0.59-0.42	0.17	7.7-4.9
Clay, Cambridge, Mass ...	1.18-0.77	0.54-0.44	0.10	6.7-5.2
Sandy clay, S. Carolina ..	1.06-0.68	0.52-0.45	0.07	6.3-5.3
Sandy loam, Va.	0.88-0.59	0.47-0.37	0.10	5.5-4.4
Sandy loam, Ill.	0.62-0.48	0.38-0.33	0.05	4.5-3.9

We see the variations of the dielectric constant $\bar{\epsilon}$, due to the settlement of clay under the pressure of a building, are rather great.

The measurement of the dielectric constant, therefore, furnishes a possibility of testing the theory of Terzaghi and, in the case of confirmation, a method for the determination of settlement.

In the case of humid rock, the dielectric constant $\bar{\epsilon}$ approaches the limiting value ϵ_c of the consolidated substance from above. On the contrary, in dry regions where, in general, the pores of rock contain little water, the dielectric constant ϵ increases in approaching the limiting value ϵ_c . In the experiments I have made with Dr. M. M. Ghali, we

* Redlich, Terzaghi und Kampe, *Ingenieur-Geologie* (Verlag Springer, Wien, 1929) fig. 222, p. 324.

LII. *On Dixon's Formula for well-poised Series.*

By T. M. MACROBERT, Professor of Mathematics
in the University of Glasgow *.

[Received November 23, 1938.]

FORMULÆ for Hypergeometric Functions and Generalized Hypergeometric Functions with unity as argument are usually particular cases or limiting cases of formulæ in which the argument is not unity. It is here shown that Dixon's Theorem (Proc. Lond. Math. Soc. ser. 1, xxxv. pp. 284-289 (1902); see also Watson, Proc. Lond. Math. Soc. ser. 2, xxii. pp. xxxii-xxxiii (1923))

$${}_3F_2\left(\begin{matrix} \alpha, \beta, \gamma; 1 \\ \alpha-\beta+1, \alpha-\gamma+1 \end{matrix}\right) = \frac{\Gamma(1+\frac{1}{2}\alpha)\Gamma(1+\alpha-\beta)\Gamma(1+\alpha-\gamma)\Gamma(1+\frac{1}{2}\alpha-\beta-\gamma)}{\Gamma(1+\alpha)\Gamma(1+\frac{1}{2}\alpha-\beta)\Gamma(1+\frac{1}{2}\alpha-\gamma)\Gamma(1+\alpha-\beta-\gamma)}, \quad (1)$$

where $R(2+\alpha-2\beta-2\gamma) > 0$, is a particular case of a more general formula, namely (4) below.

The following formula (Thomae, *Math. Ann.* ii. pp. 427-444 (1870); see also Phil. Mag. ser. 7, xxv. pp. 848-851 (1938))

$$\sum_{\alpha, \beta, \gamma} \frac{\Gamma(\beta-\alpha)\Gamma(\gamma-\alpha)}{\Gamma(\rho-\alpha)\Gamma(\sigma-\alpha)} \Gamma(\alpha) z^\alpha {}_3F_2\left(\begin{matrix} \alpha, \alpha-\rho+1, \alpha-\sigma+1; -z \\ \alpha-\beta+1, \alpha-\gamma+1 \end{matrix}\right) \\ (=) \frac{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}{\Gamma(\rho)\Gamma(\sigma)} {}_3F_2\left(\begin{matrix} \alpha, \beta, \gamma; -1/z \\ \rho, \sigma \end{matrix}\right), \quad (2)$$

where $-\pi \leq \arg z \leq \pi$, and a cross-cut is taken along the real axis from the origin to $-\infty$ to make the function uniform, gives the analytical continuation of the left-hand side, valid in the region $|z| < 1$, into the region $|z| > 1$. The symbol $(=)$ is employed to indicate that the expressions on the two sides represent the same function, each in its own domain.

To Whipple is due the identity (Proc. Lond. Math. Soc. xxvi. p. 267 (1926); see also Phil. Mag. ser. 7, xxvi. pp. 87-89 (1938))

$${}_3F_2\left(\begin{matrix} \alpha, \beta, \gamma; z \\ \alpha-\beta+1, \alpha-\gamma+1 \end{matrix}\right) = (1-z)^{-\alpha} {}_3F_2\left\{\begin{matrix} \frac{\alpha}{2}, \frac{1+\alpha}{2}, 1+\alpha-\beta-\gamma; \frac{-4z}{(1-z)^2} \\ \alpha-\beta+1, \alpha-\gamma+1 \end{matrix}\right\}, \quad (3)$$

* Communicated by the Author.

where z lies in the inner loop of the curve

$$16(x^2+y^2)=\{(x-1)^2+y^2\}^2.$$

On applying (2) to the right-hand side of (3) we obtain the formula

$$\begin{aligned} {}_3F_2\left(\begin{matrix} \alpha, \beta, \gamma; z \\ \alpha-\beta+1, \alpha-\gamma+1 \end{matrix}\right) \\ = (1-z)^{-\alpha} \frac{\Gamma(\alpha-\beta+1)\Gamma(\alpha-\gamma+1)}{\Gamma\left(\frac{\alpha}{2}\right)\Gamma\left(\frac{1+\alpha}{2}\right)\Gamma(1+\alpha-\beta-\gamma)} E, \quad (4a) \end{aligned}$$

where

$$\begin{aligned} E = & \frac{\Gamma(\frac{1}{2})\Gamma\left(1+\frac{\alpha}{2}-\beta-\gamma\right)}{\Gamma\left(\frac{\alpha}{2}-\beta+1\right)\Gamma\left(\frac{\alpha}{2}-\gamma+1\right)} \Gamma\left(\frac{\alpha}{2}\right) \left\{ \frac{(1-z)^2}{4z} \right\}^{\frac{\alpha}{2}} \\ & \times {}_3F_2\left\{ \begin{matrix} \frac{\alpha}{2}, \beta-\frac{\alpha}{2}, \gamma-\frac{\alpha}{2}; -\frac{(1-z)^2}{4z} \\ \frac{1}{2}, \beta+\gamma-\frac{1}{2}\alpha \end{matrix} \right\} \\ & + \frac{\Gamma(-\frac{1}{2})\Gamma\left(\frac{1+\alpha}{2}-\beta-\gamma\right)}{\Gamma\left(\frac{1+\alpha}{2}-\beta\right)\Gamma\left(\frac{1+\alpha}{2}-\gamma\right)} \Gamma\left(\frac{1+\alpha}{2}\right) \left\{ \frac{(1-z)^2}{4z} \right\}^{\frac{1+\alpha}{2}} \\ & \times {}_3F_2\left\{ \begin{matrix} \frac{1+\alpha}{2}, \beta+\frac{1-\alpha}{2}, \gamma+\frac{1-\alpha}{2}; -\frac{(1-z)^2}{4z} \\ \frac{3}{2}, \beta+\gamma+\frac{1-\alpha}{2} \end{matrix} \right\} \\ & + \frac{\Gamma\left(\beta+\gamma-\frac{\alpha}{2}-1\right)\Gamma\left(\beta+\gamma-\frac{1+\alpha}{2}\right)}{\Gamma(\beta)\Gamma(\gamma)} \Gamma(1+\alpha-\beta-\gamma) \\ & \times \left\{ \frac{(1-z)^2}{4z} \right\}^{1+\alpha-\beta-\gamma} {}_3F_2\left\{ \begin{matrix} 1+\alpha-\beta-\gamma, 1-\beta, 1-\gamma; -\frac{(1-z)^2}{4z} \\ 2+\frac{\alpha}{2}-\beta-\gamma, \frac{3+\alpha}{2}-\beta-\gamma \end{matrix} \right\}, \quad (4b) \end{aligned}$$

and z lies in the region between the two loops of the curve. This region contains the point $z=1$.

If now $R(2+\alpha-2\beta-2\gamma)>0$, Dixon's formula (1) is obtained by putting $z=1$.

On putting $\frac{1}{2}(\alpha+1)$ for γ in (4), or by applying analytical continuation to Gauss's formula,

$$F\left(\alpha, \beta; z\right) = (1-z)^{-\alpha} F\left\{\begin{matrix} \frac{1}{2}\alpha, \frac{1}{2} + \frac{1}{2}\alpha - \beta; \\ 1 + \alpha - \beta \end{matrix}; \frac{-4z}{(1-z)^2}\right\}, \quad (5)$$

we obtain the corresponding formula

$$\begin{aligned} F\left(\alpha, \beta; z\right) &= \frac{\Gamma(\alpha-\beta+1)\Gamma(1-2\beta)}{\Gamma(1-\beta)\Gamma(1+\alpha-2\beta)} z^{-\frac{1}{2}\alpha} F\left\{\begin{matrix} \frac{\alpha}{2}, \beta - \frac{\alpha}{2}; \\ \beta + \frac{1}{2} \end{matrix}; -\frac{(1-z)^2}{4z}\right\} \\ &+ \frac{\Gamma(\alpha-\beta+1)\Gamma(2\beta-1)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(1-z)^{1-2\beta}}{z^{\frac{1}{2} + \frac{1}{2}\alpha - \beta}} \\ &\times F\left\{\begin{matrix} \frac{1-\alpha}{2}, \frac{1+\alpha}{2} - \beta; \\ \frac{3}{2} - \beta \end{matrix}; -\frac{(1-z)^2}{4z}\right\}. \quad (6) \end{aligned}$$

Note.—The function whose values in the regions $|z| < 1$, $|z| > 1$, are given by formula (2) is uniform over the entire plane, bounded by the cross-cut. If $R(\rho + \sigma - \alpha - \beta - \gamma) > 0$ the series on both sides converge absolutely for points on the circle $|z| = 1$, and consequently at such points the two sides can be equated. In particular, on putting $z = \exp(\pm i\pi)$ a formula employed by Whipple (Proc. Camb. Phil. Soc. xxi. p. 501 (1923)) in proving a result due to Thomae and Ramanujan is obtained. Similar formulæ can, of course, be found in the same way for functions ${}_{p+1}F_p$, where p is greater than 2.

LIII. *Thermal Instability of Dielectrics for alternating Voltages
when the Loss Angle is dependent upon the Field Strength.*

By A. GEMANT, University of Wisconsin, U.S.A.,
and S. WHITEHEAD *†.

[Received January 6, 1939.]

SUMMARY.

The theory of the thermal instability of dielectrics is extended for alternating voltages with a conductivity dependent upon the field strength. Calculations are made with the simplifying assumption that the field is determined by the distribution of the dielectric constant alone and that the variation with temperature of the latter can be neglected. The equations of thermal stability and of their limiting cases are developed for three main types of the conductivity-temperature function, and formulæ for the breakdown voltage are thence derived for several types of conductivity-field strength function. These formulæ are illustrated by application to some experimental results, showing the variation of breakdown voltage which may be anticipated from certain materials tested. Certain limitations and extensions required in practice are discussed.

I. *General.*

THE theory of the thermal breakdown of dielectrics, the principles of which were first laid down by Miles Walker, Everest, and K. W. Wagner⁽¹⁾, was developed further mainly through the work of V. Fock⁽²⁾, W. Rogowski, and Dreyfus⁽³⁾. In recent years research has continued on more experimental lines. From the large number of publications only a few need be mentioned: J. B. Whitehead⁽⁴⁾ developed life tests for impregnated paper insulation, S. Whitehead and W. Nethercot⁽⁵⁾ investigated the instability of dielectrics in relation to their electrical conductivity, two papers by D. R. Robinson⁽⁶⁾ and by L. G. Brazier⁽⁷⁾ are concerned with the breakdown of cables, P. Perlick⁽⁸⁾, on the one hand, K. W. Wagner and A. Gemant⁽⁹⁾ on the other, investigated the variations with frequency of the thermal electric strength.

As a result of these investigations and a number of others carried out

* Communicated by E. B. Wedmore.

† The paper is based on Report Ref. L/T88 of the British Electrical and Allied Industries Research Association.

under the E. R. A.—those by L. Hartshorn and E. Rushton⁽¹⁰⁾ and L. Hartshorn and W. H. Ward⁽¹¹⁾ should be mentioned here—there has arisen the need for a further development of the theory of thermal breakdown to cases where the electrical conductivity is a function of the field strength. The investigations quoted above led to the result that the conductivity depends upon the field strength in a great number of cases such, it is to be expected, as to affect the electric strength. It is also known that such variations with electric stress may occur in insulation which has deteriorated owing to ingress of moisture or other causes and which may eventually fail due to thermal instability.

The formulæ of Fock, originally developed for D.C., can be used, but become very involved in a generalized treatment for D.C. except in two particular cases, where the conductivity is proportional to the field strength. With A.C. it is fortunately possible to assume that the field strength is substantially constant, which facilitates solution. In addition, practical applications are more frequent with A.C., so that the present study mainly applies to alternating fields, though it is to be anticipated that the D.C. problem will become equally important in the future.

The assumption that the field strength is constant implies that the quantity $\epsilon\sqrt{(1+\tan^2 \delta)}$ is approximately constant, where ϵ is the dielectric constant and δ the loss angle. Since, except in special cases, $\tan \delta$ rarely exceeds 0.1, its variation is unimportant. If κ is the polarizability of the dielectric, then

$$\kappa = \frac{3}{4\pi} \frac{\epsilon - 1}{\epsilon + 2} \quad \dots \dots \dots (1)$$

and $d\epsilon/\epsilon : d\kappa/\kappa = (\epsilon + 2)(\epsilon - 1)/3\epsilon \dots \dots \dots (2)$

If t is the temperature in degrees centigrade $(1/\kappa)(d\kappa/dt)$ is minus three times the linear thermal expansion, and so is of the order of -20.10^{-6} , which will also be the order of $(1/\epsilon)(d\epsilon/dt)$ from this cause, since ϵ is of the general order of 3. Even for the new high permittivity materials such as Condensa, Kerapur etc., where ϵ may be 60, $(1/\epsilon)(d\epsilon/dt)$ is about 400.10^{-6} , and this effect then determines the temperature coefficient.

So far only the linear expansion has been considered, which is outweighed for dielectrics of normal permittivity by absorption, dipolar, and other polarization mechanisms. The order of the temperature coefficient from this cause is about 10^{-4} , *e. g.*, $+5.10^{-4}$ for ebonite, $+10^{-4}$ for mica, $+5.10^{-4}$ for porcelain etc.

In this way the total variation of $\epsilon\sqrt{(1+\tan^2 \delta)}$ for a temperature change of, say, 100°C. will not exceed about 5 per cent., and the field strength E may therefore be assumed constant to the same accuracy*.

* A formula for correction for the temperature coefficient of ϵ was given by Dreyfus (*loc. cit.*), and in practice this correction has always been found negligible.

II. Derivation of the General Equations.

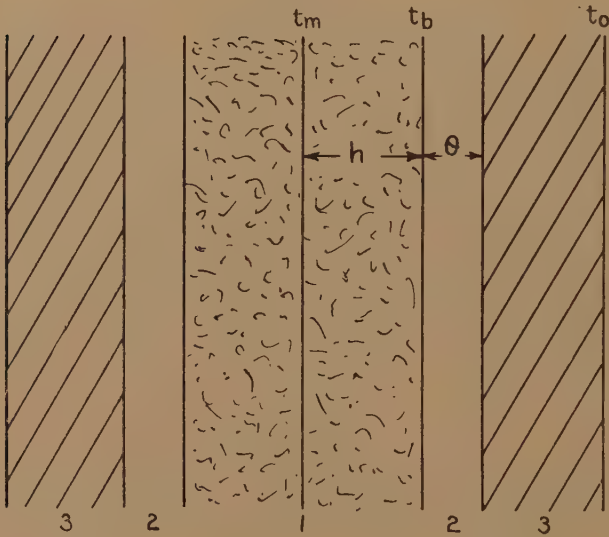
Let the equivalent A.C. conductivity σ be given by

$$\sigma = \sigma_0 f(E) \cdot E^2 e^{at}, \quad \dots \dots \dots (3)$$

where σ_0 and a =constants, t =temperature in deg. C., and $f(E)$ =any function of the field strength. Two other possibilities which are not covered by equation (3) will be dealt with in the later sections. We note that

$$\sigma = \frac{1}{2} \nu \epsilon \tan \delta, \quad \dots \dots \dots (4)$$

Fig. 1.



Cross-section through insulator.

where ν =frequency ; it can be computed therefore from measurements of ϵ and $\tan \delta$; at 50 cyc./sec.

$$\sigma = 0.278 \times 10^{-10} \epsilon \tan \delta \text{ watts/cm.}^3 \text{ per \{volt (r.m.s.)/cm.\}^2.}$$

A parallel field is considered, shown diagrammatically in fig. 1, where 1 is the dielectric of thickness $2h$, 2 are the electrodes each of thickness θ , and 3 are the packing layers which may or may not be present. The axis parallel to the field is designated by z , its zero point lying in the plane of symmetry. The temperature in this latter is denoted by t_m , the temperature at the boundary of dielectric and metal by t_b ; the ambient temperature is t_0 .

The solution for a form such as equation (3), with the simplifying

Eliminating t_b from equations (12) and (13),

$$k \sqrt{\left(\frac{2\beta y}{a}\right) \tanh \eta} = \frac{k_m}{a\theta} \log \{y(1 - \tanh^2 \eta)\} - \frac{k_m t_0}{\theta}, \quad (14)$$

where

$$\eta = h \sqrt{\frac{1}{2} \alpha \beta y}. \quad (15)$$

Equation (14) gives the equilibrium temperature (y) corresponding to a given field strength (β).

The breakdown voltage is the maximum value which permits a real solution to equation (14). It may be noted that, owing to the assumption of a continuous field strength, the corresponding limiting temperatures are finite, whereas this, as has previously been shown, is not always the case with D.C. where the current is continuous. It is convenient to introduce the parameter c , which is the ratio of the internal to the external thermal resistance; thus

$$c = P_i/P_e = k_m h / k \theta, \quad (16)$$

giving $\log y = (2/c) \eta \tanh \eta + \log \cosh^2 \eta + a t_0. \quad (17)$

The limiting condition just mentioned corresponds to the vanishing of the Jacobian function, which in this case corresponds to the maximum condition for V , the vanishing of dV/dt_m or $d\beta/d\eta$, since these give only two variables, thus:

$$-\frac{1}{2\beta} \frac{d\beta}{d\eta} = \frac{1}{2} \frac{d \log y}{d\eta} - \frac{1}{\eta} = \left(\frac{1}{c} + 1\right) \frac{\tanh \eta}{\eta} + \frac{1}{c \cosh^2 \eta} - \frac{1}{\eta^2} = 0. \quad (18)$$

For thin layers c is small, and from (15) η becomes $\sqrt{\frac{1}{2}c}$, so that

$$\log y = 1 + a t_0 \quad (19)$$

and

$$\beta = c e^{-a t_0} / a e h^2, \quad (20)$$

which gives the breakdown voltage V if $f(E)$ in (7) is known, since

$$\sigma_0 f(E) V^2 = h e^{-a t_0} / a e P_e. \quad (21)$$

If, for example, $f(E)$ is E , then

$$V = (8 k h c / e a \sigma_0 e^{a t_0})^{1/3} \text{ e.s.u.}, \quad (22)$$

or, for A.C. at 50 cyc./sec.,

$$E = (360 / e P_e a h \epsilon \tan \delta)^{1/3} \text{ kv. (r.m.s.) / mm.}, \quad (23)$$

the form given by Whitehead and Nethercot, if P_e is in thermal ohms, *i. e.*, deg. cent. per watt per cm.² of surface, and δ the loss angle at a stress of 1 kv./mm.

On the other hand, if the specimen is very thick or if the external thermal resistance is small c is great, and (18) becomes

$$\eta \tanh \eta = 1, \quad (24)$$

whence η is 1.20 and $\log y$ is $(1.187 + at_0)$, so that

$$\beta = 0.88e^{-at_0}/ah^2. \quad (25)$$

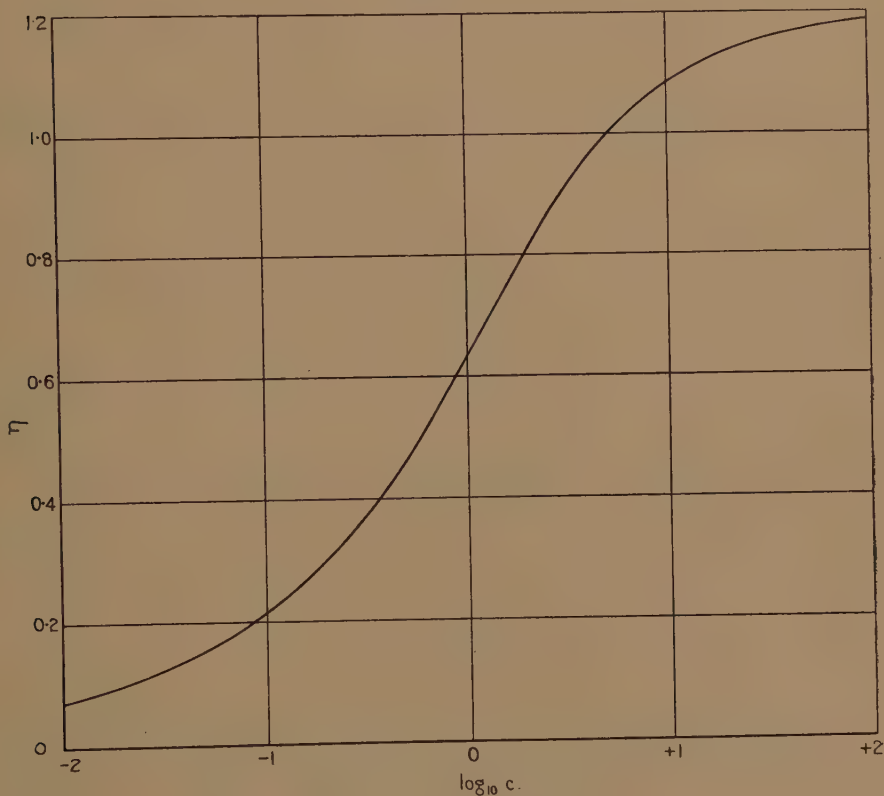
If $f(E)$ is unity

$$V = 1.88(k/a\sigma_0 e^{at_0})^{1/2}, \quad (26)$$

and is independent of thickness. If $f(E)$ is E

$$V = 1.92(kh/a\sigma_0 e^{at_0})^{1/3}. \quad (27)$$

Fig. 2.



η as function of $\log c$.

Thick specimens have already been studied by Dreyfus and Berger. For intermediate thicknesses fig. 2 gives the value of η from (18) for values of c from 0.01 to 100. Equation (17) is now written

$$\log y = F(c) + at_0, \quad (28)$$

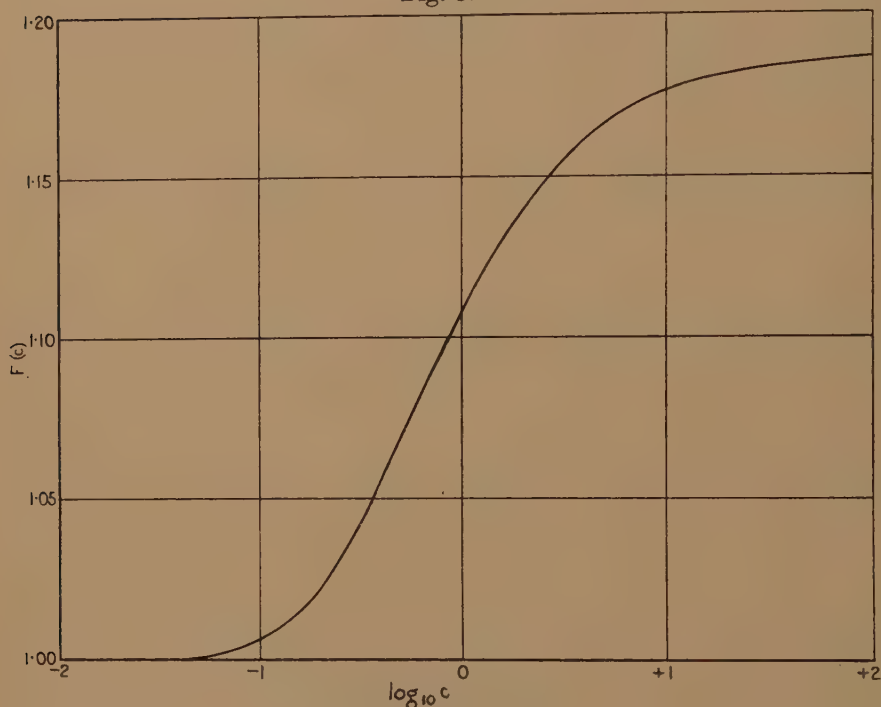
the solution of which is plotted in fig. 3. Thence β may be evaluated from (15) and V from (7). If V_m is the breakdown voltage for a thin

specimen then, from (21),

$$\left. \begin{aligned} V^2 f(E)/V_m^2 f(E_m) &= 2kPe\eta^2/hye^{-at^0} \\ &= 2e\eta^2/cye^{-at^0} \\ &= 2e\eta^2/ce^{F(c)} \end{aligned} \right\} \dots \dots (29)$$

$$= \text{function of } c \text{ only} \dots \dots (30)$$

Fig. 3.



$F(c)$ as function of $\log c$.

Equation (30) illustrates the effect of thermal resistance in the dielectric, showing at what value of c the assumption of a thin dielectric is no longer true. The right-hand side of the equation is the same function as that evaluated by Whitehead and Nethercot in fig. 1 of their paper. As was mentioned by them, the closeness of the approximation of this function to unity measures the degree of approximation of V to V_m whatever may be the dependence of the conductivity on field strength.

III. Special Conductivity Functions.

(a) Conductivity Independent of Field Strength.

This reduces simply to

$$V = \{8k\eta^2/a\sigma_0 y\}^{1/2} \dots \dots \dots (31)$$

(b) *Linear Function of Field Strength.*

Let $\sigma = \sigma_0(1 + bE)e^{at}$.
 Then
$$V = (8k\eta^2/a\sigma_0y)^{1/2} \left(1 + \frac{bV}{2h}\right)^{-1/2} \dots \dots \dots (32)$$

$$= \{16kh\eta^2/a\sigma_0by\}^{1/3}$$

if $bV/2h$ is large, as is often the case.

(c) *Exponential Function of Field Strength.*

Let $\sigma = \sigma_0 e^{\alpha E} e^{at}$.
 Then from (15)
$$Ve^{\alpha V/4h} = \{8k\eta^2/a\sigma_0y\}^{1/2}, \dots \dots \dots (33)$$

which can be solved by trial easily if V is known approximately. This case arises frequently when ionization occurs in dielectrics, and may also occur due to mobility changes owing to the existence of a complex viscosity in elasto-viscous dielectrics⁽¹⁴⁾. If α is small this case tends to become case (b).

(d) *Conductivity containing a Constant Term.*

Let $\sigma = p + \sigma_0^f(E)e^{at} \dots \dots \dots (34)$

In the neighbourhood of room-temperature the constant term is frequently the dominating one, but this evidently means that the thermal breakdown voltage is so high that it does not come into play, since it will be preceded by electric breakdown or breakdown due to chemical decomposition at the high but theoretically stable internal temperatures. In consequence of this true thermal instability is confined mainly to higher ambient temperatures, and for present purposes the second term is taken as dominant, and p is introduced merely as a correction.

Equation (5) now becomes

$$d^2t/dz^2 = -p' - \beta e^{at},$$

where
$$p' = pE^2/k. \dots \dots \dots (35)$$

Instead of equation (8) we have

$$z = - \int \sqrt{\left\{\frac{2\beta}{a}\right\} \left\{(e^{at_m} - e^{at}) + \frac{ap'}{\beta}(t_m - t)\right\}} dt \dots \dots \dots (36)$$

This equation is still correct, but is not integrable in finite form. The second term below the root is therefore considered as small compared with the first, and the root expanded as a series to the second term, giving

$$z = - \int \sqrt{\frac{2\beta}{a}(e^{at_m} - e^{at})} dt + p' \left(\frac{a}{2\beta y}\right)^{3/2} \int \frac{(t_m - t)dt}{\{1 - e^{a(t_m - t)}\}^{3/2}} \dots \dots (37)$$

The exponential is then also expanded in the second term. This procedure would be unjustified in the main term, but permissible in the correction term. The integration now gives

$$z = \sqrt{\frac{2}{a\beta y}} \tanh^{-1} \sqrt{\frac{y - e^{at}}{y}} - \frac{2p'\sqrt{(t_m - t)}}{(2\beta y)^{3/2}}. \quad (38)$$

But when $v < 1$, $\tanh(u + v) \simeq \tanh u + v \operatorname{sech}^2 u$, so that the following may be obtained :

$$t = \frac{1}{a} \log \{y[1 - \tanh^2 z \sqrt{2\beta y/a}]\} - \frac{p'}{\sqrt{2\beta y}} z \sqrt{(t_m - t)}. \quad (39)$$

Approximately, from (8) $\sqrt{(t_m - t)} \simeq z \sqrt{\beta y/2}$, which, when substituted in (39), gives

$$t = (1/a) \log \{y[1 - \tanh^2 z \sqrt{2\beta y/a}]\} - \frac{1}{2} p' z^2. \quad (40)$$

and
$$dt/dz = -\sqrt{2\beta y/a} \tanh \{z \sqrt{\alpha \beta y/2}\} - p' z. \quad (41)$$

Equations (40) and (41) represent the first approximations to the solution, together with the boundary conditions

$$t = t_b \text{ and } dt/dz = k_m(t_b - t_0)/\theta, \text{ when } z = h. \quad (42)$$

Equation (19) also becomes

$$\log y = (2/c)\eta \tanh \eta + 2 \log \cosh \eta + p' a h^2 \{c^{-1} + 2^{-1}\} + a t_0. \quad (43)$$

Equation (18) remains unchanged, and with (43) gives the breakdown strength.

Let $f(E)$ be E , then

$$y = 16k\eta^2 h / a \sigma_0 V^3,$$

and (43) becomes

$$3 \log V + \frac{pa}{4k} \left(\frac{1}{c} + \frac{1}{2} \right) V^2 = \log \left(\frac{16k\eta^2 h}{a \sigma_0} \right) - F(c) - a t_0, \quad (44)$$

which may now be solved by trial. It is to be remarked that the constant term lowers the breakdown voltage.

(e) Exponential Conductivity Function limited to Higher Temperatures.

It is of interest also to discuss the case when σ is not a simple exponential function of temperature. At lower temperatures it is either roughly constant or passes through a maximum before beginning to increase exponentially. This occurs early if the material is moist or otherwise poor, but the temperature at which it is pronounced increases with increasing quality of the material. It is possible to obtain a good approximation to the real behaviour by assuming that $\tan \delta$ is *independent*

of the temperature up to a certain temperature t_x , and that it is given by equation (3) above this temperature. Thus

$$\sigma = \sigma_0 f(E) e^{at}, \quad t \geq t_x, \quad (45)$$

$$\sigma = \sigma_0 f(E) e^{at_x}, \quad t < t_x. \quad (46)$$

Equations (45) and (46) mean that the dielectric sheet is made up of two layers, one of which "a" obeys (45), the other "b" equation (46). Since the temperature decreases from the plane of symmetry, layer "a" is situated in the middle, surrounded at both sides by two sheets "b." The respective thickness of the layers varies with the field strength, since the characteristic of the boundary is its temperature t_x . The thickness of the middle layer, $2z_x$, is one of the variables.

The solution for layer "a" is given in Section II., while that for layer "b" is easily shown to be

$$t = t_x + \frac{1}{2} q (z_x^2 - z^2) + B(z_x - z), \quad (47)$$

$$q = \sigma_0 f(E) e^{at_x} \times E^2/k. \quad (48)$$

By evaluating the flux of heat at the outer boundary of the dielectric

$$B = (c/h)(t_b - t_0) - qh, \quad (49)$$

whence
$$t_b = \frac{t_x - \{(c/h)t_0 + qh\}(z_x - h) + \frac{1}{2}q(z_x^2 - h^2)}{1 - (c/h)(z_x - h)} \quad (50)$$

By inserting the values of B and t_b in the equations for continuity of t and dt/dz at the boundary between the layers the following equations are obtained:

$$\sqrt{2/\alpha\beta y} \tanh^{-1} \sqrt{\{(y - y_x)/y^2\}} - z_x = 0, \quad (51)$$

$$\frac{1}{y} \sqrt{\left\{ \frac{2(y - y_x)}{\alpha\beta} \right\}} - \frac{(1/\beta y_x)(t_x - t_0) + (h/c)(z_x - h) - \frac{1}{2}(z_x - h)^2}{(h/c) - (z_x - h)} = 0, \quad (52)$$

where
$$y_x = e^{at_x} = q/\beta. \quad (53)$$

Now the condition for the limit of thermal stability is the limiting condition for a real solution of equations (51) and (52), namely, the vanishing of the Jacobian function of the two equations, that is

$$\sqrt{\beta} \left(\frac{h}{c} - z_x + h \right) = \sqrt{\left\{ \frac{2(y - y_x)}{\alpha y_x^2} \right\}} / \left\{ \frac{(y/y_x)}{1 - \sqrt{\{(y - y_x)/y\}} \tanh^{-1} \sqrt{\{(y - y_x)/y\}}} - 1 \right\}} \quad (54)$$

Equations (51), (52), and (54) permit the computation of β and thence

the breakdown voltage. The convenient method is to assume a value for y and to calculate $z_x\sqrt{\beta}$ from (51) and $\{(h/c)-z_x+h\}$ from (54), when (52) may be checked and a better value for y chosen.

(f) *Temperature Exponent Linear Function of Field Strength.*

In some cases it was found that the exponent, a , itself is a linear function of field strength. In dealing with this case let us combine it with a constant term in p and write

$$\sigma = p + \sigma_0 E e^{(a - \lambda E)t} \quad . \quad . \quad . \quad . \quad . \quad (55)$$

Here equation (44) may be used in replacing a by $a - \lambda E$. The logarithmic term of the right-hand side can be developed in a series to the second term, since $\lambda E < a$. In doing so the following equation is finally deduced :

$$3 \log V - \frac{\lambda}{2h} \left(\frac{1}{a} + t_0 \right) V + \frac{pa}{4k} \left(\frac{1}{c} + \frac{1}{2} \right) V^2 \\ - \frac{p\lambda}{8kh} \left(\frac{1}{c} + \frac{1}{2} \right) V^3 = -\log \frac{16 k \eta^2 h}{a \sigma_0} - F(c) - at_0. \quad . \quad . \quad . \quad (56)$$

IV. Numerical Examples.

Whitehead and Nethercot* have given data on various materials in different conditions which show the features considered in the foregoing. Section III. (a) applies to the results on a special thin ebonite and a dry condenser tissue where the conductivity was independent of field strength, while Section III. (b) applies to cellulose acetate and cellulose ethyl ether where the conductivity may be taken as a linear function of the field. In point of fact the best approximation in the latter cases is to take

$$\sigma = \sigma_0 b E e^{at},$$

so that

$$V = \left\{ \frac{16 k h \eta^2}{a \sigma_0 b y} \right\}^{1/3} = 2 \left\{ \frac{2 h \eta^2}{a u} \cdot \frac{3.6 \times 10^{10}}{T \epsilon \tan \delta} \right\}^{1/3} \text{ volts at 50 cyc./sec.} \quad . \quad (57)$$

where $T = 1/k =$ thermal resistivity of dielectric in thermal ohm-cm., $u = e^{F(c)}$, and $\epsilon \tan \delta = \sigma_0 b e^{at_0} =$ energy losses at ambient temperature and 1 volt/cm. field strength.

Section III. (e) applies to dried cellulose acetate where the temperature function is limited to the higher temperatures. For dry grease-proof paper an exponential variation of conductivity with field strength is appropriate as in Section III. (c), but the equation is expressed as

$$V e^{aV/4h} = 2 \left(\frac{2 \eta^2}{a u} \cdot \frac{3.6 \times 10^{10}}{T \epsilon \tan \delta} \right) \text{ at 50 cyc./sec.,} \quad . \quad . \quad (58)$$

* Reference (5).

where u and T have the meanings assigned and δ =loss angle when $\alpha E < 1$.

The breakdown voltages so computed * are plotted in fig. 4 as functions of c . The parameter c is the ratio P_i/P_e , but in some instances it is necessary to know the thickness. Accordingly it has been assumed that P_e is fixed and is such that $c=1$ when $h=1$ or the thickness is 2 cm.—that is, P_e is taken as numerically equal to the thermal resistivity of the material. On this basis c may also be taken numerically as one-half the thickness of the dielectric in cm., while (57) becomes

$$V = 2\{(2c\eta^2/au)(3.6 \times 10^{10}/T\epsilon \tan \delta)\}^{1/3} \text{ volts at 50 cyc./sec.} \quad (59)$$

Fig. 4 shows, as already indicated, that variation of conductivity with stress reduces the breakdown voltage at small thicknesses, since the electric strength increases less rapidly as the thickness decreases. At great thicknesses, where the limiting voltage is approached, variation of conductivity with field strength is not so important. In fact fig. 4 is inaccurate for large values of c for those instances where a field strength function is assumed, since the constant term has been neglected, but is actually important in these circumstances since the electric strength tends to zero at great thicknesses. On this account the cellulose materials would not have so high a limiting voltage as indicated in fig. 4; in actual fact they would be of the same order or less than those for ebonite and dry condenser tissue.

Certain limitations are required to the theory in practice. It is necessary, for its validity, that the critical or decomposition temperature t_c should not be exceeded. This condition is

$$(2/c)\eta \tanh \eta + \log \cosh^2 \eta < a(t_c - t_0). \quad (60 a)$$

If the dielectric is very thick, or if the surface is maintained at the ambient temperature, this condition reduces to

$$1.187 < a(t_c - t_0). \quad (60 b)$$

If the dielectric is very thin the condition becomes

$$\sqrt{2/c} \tanh \sqrt{c/2} + \log \cosh^2 \sqrt{c/2} < a(t_c - t_0), \quad (60 c)$$

or $1/a < t_c$ in the limit †.

These conditions do not take into account the constant term denoted by p earlier in the text, which, for thick dielectrics, frequently causes a

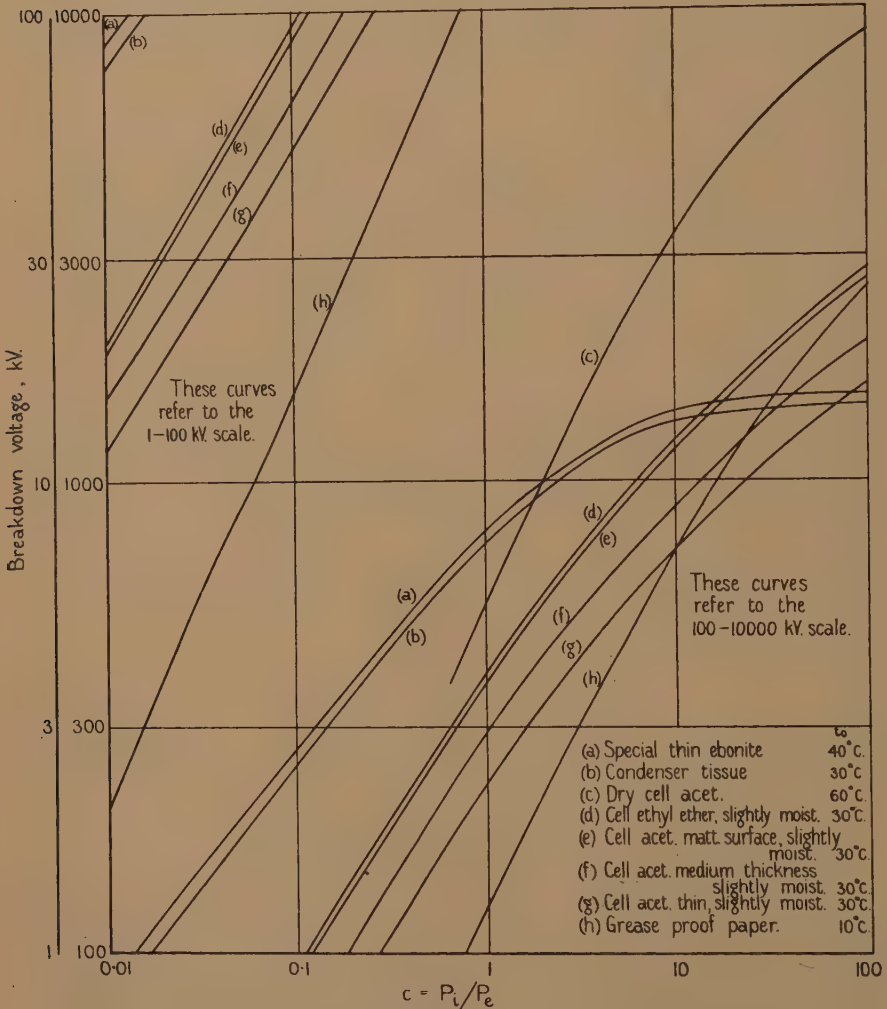
* Numerical values are taken from those given in reference (5).

† In some cases it is therefore impossible for a thin dielectric, whatever the thermal conditions applied, to become unstable in the manner envisaged. In practice, however, one does not distinguish between true thermal instability and breakdown due to a steady rise to the critical temperature.

considerable increase in the temperature. Where, in fact, the constant term is dominant with a very thick dielectric we have, as indicated elsewhere⁽¹⁵⁾,

$$V = [8k(t_c - t_0)/p]^{1/2} \dots \dots \dots (61)$$

Fig. 4.



Variation of breakdown voltage with parameter c (i.e., thickness in case considered) for various materials.

It is therefore necessary to ensure in a given instance that V from equation (61) is considerably greater than the V calculated by the expressions given earlier in the paper.

Nevertheless fig. 4 gives legitimate results over the range most important in practice, and in general would only need important corrections for values of c greater than 10. In any case it shows the main features which arise from the present analysis.

References.

- (1) Miles Walker, "Theory of Machine Design." Also, J. I. E. E. xlix. p. 3, Discussion. Everest, Discussion on Paper by Evershed, J. I. E. E. lii. p. 51 (1913). K. W. Wagner, J. A. I. E. E. xli. p. 1034 (1922).
- (2) V. Fock, *Arch. f. Elektrotechn.* xix. p. 71 (1927).
- (3) W. Rogowski, *Arch. f. Elektrotechn.* xiii. p. 168 (1924). Dreyfus, S. E. V. Bulletin, 7 and 15 (1924). Berger, S. E. V. Bulletin (February 1926).
- (4) J. B. Whitehead, Trans. A. I. E. E. lii. p. 1004 (1933).
- (5) S. Whitehead and W. Nethercot, Proc. Phys. Soc. xlvii. p. 974 (1935).
- (6) D. R. Robinson, J. I. E. E. lxxvii. p. 90 (1935).
- (7) L. G. Brazier, J. I. E. E. lxxvii. p. 104 (1935).
- (8) P. Perlick, Dissertation Techn. Hochsch. Berlin (1934).
- (9) K. W. Wagner and A. Gemant, *Berliner Akad. Ber.* (1934).
- (10) L. Hartshorn and E. Rushton, E. R. A. Report, Ref. L/T60 (1935). In course of publication.
- (11) L. Hartshorn and W. H. Ward, E. R. A. Report, Ref. L/T48 (1932). See J. I. E. E. lxxvii. p. 467 (Nov. 1935).
- (12) E. Albers-Schönberg, *Zeits. d. V. D. I.* lxxviii. p. 892 (1934).
- (13) H. Handrek, *Arch. f. Techn. Messen*, iv. p. 28 (1935).
- (14) A. Gemant, Trans. Farad. Soc. xxxi. p. 1582 (1935).
- (15) S. Whitehead, Brit. Assoc. (1936) ; World Power, vol. xlvi. p. 327 (1936).

LIV. *Direct Determination of Stresses from the Stress Equations in some Two-Dimensional Problems of Elasticity.*—Part III. *Problems of Non-Isotropic Material.*

By BIBHUTIBHUSAN SEN, Krishnagar College, Bengal *.

[Received August 15, 1938.]

IX. *Statement of the Problem.*

THE object of this paper is to show that two-dimensional problems connected with certain bodies of non-isotropic material can be solved in as simple a manner as the corresponding problems of isotropic bodies discussed previously †. We shall consider, for example, a simple type of non-isotropic body in the form of an infinitely large plate with a horizontal straight boundary which has different moduli of elasticity along the horizontal and vertical directions. It is found that for various distribution of loads on the straight horizontal boundary, stresses can be deduced directly from the stress equations of equilibrium and compatibility.

The axis of x is taken along the straight boundary and the axis of y is drawn into the plate perpendicular to this line. Then for the plate described above, the strain energy function W is given by

$$2W = Ae_{xx}^2 + Be_{yy}^2 + 2Ce_{xx}e_{yy} + \mu e_{xy}^2, \quad \dots \quad (9.1)$$

so that the stresses are obtained as

$$\left. \begin{aligned} \widehat{xx} &= Ae_{xx} + Ce_{yy}, \\ \widehat{yy} &= Ce_{xx} + Be_{yy}, \\ \widehat{xy} &= \mu e_{xy}, \end{aligned} \right\} \dots \quad (9.2)$$

where A , B , C , and μ are constants.

Hence the strain components can be written in the forms

$$\left. \begin{aligned} e_{xx} &= \frac{1}{AB-C^2} [B\widehat{xx} - C\widehat{yy}], \\ e_{yy} &= \frac{1}{AB-C^2} [A\widehat{yy} - C\widehat{xx}], \\ e_{xy} &= \frac{1}{\mu} \widehat{xy}. \end{aligned} \right\} \dots \quad (9.3)$$

* Communicated by the Author.

† *Vide* Phil. Mag. (7) xxvi. p. 98 (1938).

Again, if

E_1, E_2 be Young's moduli corresponding to extensions along ox and oy respectively, and

σ_1, σ_2 , the Poisson's ratios for contractions along these directions, we can put

$$\left. \begin{aligned} e_{xx} &= \frac{\widehat{xx}}{E_1} - \frac{\sigma_1}{E_2} \widehat{yy}, \\ e_{yy} &= \frac{\widehat{yy}}{E_2} - \frac{\sigma_2}{E_1} \widehat{xx}, \\ e_{xy} &= \frac{1}{\mu} \widehat{xy}. \end{aligned} \right\} \dots \dots \dots (9.4)$$

On comparing (9.4) with (9.3) we get

$$\frac{\sigma_1}{E_2} = \frac{\sigma_2}{E_1} = L \text{ (say)}. \dots \dots \dots (9.5)$$

Considering the deformation of a rectangular parallelepiped in which $\widehat{yy} = -\widehat{xx}$ and $\widehat{zz} = 0$, we get to the first order of approximation

$$\frac{1}{\mu} = L \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} + 2 \right). \dots \dots \dots (9.6)$$

The body stress equations of equilibrium are

$$\left. \begin{aligned} \frac{\partial \widehat{xx}}{\partial x} + \frac{\partial \widehat{xy}}{\partial y} &= 0, \\ \frac{\partial \widehat{xy}}{\partial x} + \frac{\partial \widehat{yy}}{\partial y} &= 0, \end{aligned} \right\} \dots \dots \dots (9.7)$$

and the equation of compatibility satisfied by the strain components is

$$\frac{\partial^2 e_{xx}}{\partial y^2} + \frac{\partial^2 e_{yy}}{\partial x^2} = \frac{\partial^2 e_{xy}}{\partial x \partial y}. \dots \dots \dots (9.8)$$

Substituting the results (9.4) in (9.8) we obtain

$$\frac{\partial^2}{\partial y^2} \left[\frac{\widehat{xx}}{\sigma_2} - \widehat{yy} \right] + \frac{\partial^2}{\partial x^2} \left[\frac{\widehat{yy}}{\sigma_1} - \widehat{xx} \right] = \frac{\partial^2}{\partial x \partial y} \left[\left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} + 2 \right) \widehat{xy} \right],$$

from which we get with the help of the equations (9.7), the result

$$\left(\frac{\partial^2}{\partial x^2} + k^2 \frac{\partial^2}{\partial y^2} \right) \odot = 0, \dots \dots \dots (9.9)$$

where

$$k^2 = \frac{\sigma_1}{\sigma_2} = \frac{E_2}{E_1} \dots \dots \dots (9.10)$$

and

$$\odot = \widehat{xx} + \widehat{yy}. \dots \dots \dots (9.11)$$

Putting

$$y = ky_1 \dots \dots \dots (9.12)$$

the equation (9.9) can be obtained in the form

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y_1^2} \right) \odot = 0. \dots \dots \dots (9.13)$$

The result (9.9) also enables us to write the equations of equilibrium as

$$\left. \begin{aligned} \nabla_1^2 \widehat{xx} + \frac{1}{k^2} \frac{\partial^2 \odot}{\partial x^2} &= 0, \\ \nabla_1^2 \widehat{yy} + k^2 \frac{\partial^2 \odot}{\partial y^2} &= 0, \\ \nabla_1 \widehat{xy} + \frac{\partial^2 \odot}{\partial x \partial y} &= 0, \end{aligned} \right\} \dots \dots \dots (9.14)$$

where

$$\nabla_1^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \dots \dots \dots (9.15)$$

X. Solution suitable for a Straight Boundary.

It is apparent from the equation (9.13) that we can put

$$\left. \begin{aligned} \odot &= f(z_1) \\ z_1 &= x + iy_1 = x + \frac{iy}{k} \end{aligned} \right\} \dots \dots \dots (10.1)$$

As the solution of the equations (9.14) we can write

$$\left. \begin{aligned} \widehat{yy} &= \frac{k^2}{k^2 - 1} \operatorname{Re}[f(z_1) + \phi_1(z)], \\ \widehat{xy} &= \frac{k}{1 - k^2} \operatorname{Re}[if(z_1) + \phi_2(z)], \\ \widehat{xx} &= \frac{1}{1 - k^2} \operatorname{Re}[f(z_1) + k^2 \phi_1(z)], \end{aligned} \right\} \dots \dots \dots (10.2)$$

in which z_1 has the value given in (10.1) and

$$z = x + iy.$$

In general $\phi_1(z)$, $\phi_2(z)$ and $f(z_1)$ will be completely determined by the given values of \widehat{yy} and \widehat{xy} on the boundary $y=0$, and the stress equations (9.7). For example, if the boundary $y=0$ be subjected to normal stress only, we have

$$\widehat{xy} = 0 \quad \text{when} \quad y = 0,$$

so that we can put

$$\widehat{xy} = \frac{k}{1 - k^2} \operatorname{Re}[i\{f(z_1) - f(z)\}]. \dots \dots \dots (10.3)$$

The stress equations will be satisfied if we now take

$$\widehat{y}y = \frac{k}{k^2-1} \operatorname{Re}[kf(z_1) - f(z)], \quad . \quad . \quad . \quad . \quad . \quad (10.4)$$

and

$$\widehat{x}x = \frac{k}{k^2-1} \operatorname{Re}[f(z) - \frac{1}{k}f(z_1)]. \quad . \quad . \quad . \quad . \quad . \quad (10.5)$$

On the straight boundary $y=0$, we get

$$(\widehat{y}y)_{y=0} = \frac{k}{k+1} \operatorname{Re}f(z). \quad . \quad . \quad . \quad . \quad . \quad (10.6)$$

From the above value of $\widehat{y}y$ on the straight boundary $y=0$ we get the function $f(z)$, from which $f(z_1)$ can be written down on replacing z by z_1 . Hence the stress components which involve these two functions only are completely determined.

Again, if there be a distribution of shear-load only on the boundary $y=0$, we shall have

$$\widehat{y}y = 0 \quad \text{when} \quad y=0,$$

so that we can put

$$\widehat{y}y = \frac{k^2}{k^2-1} \operatorname{Re}[\psi(z_1) - \psi(z)]. \quad . \quad . \quad . \quad . \quad . \quad (10.7)$$

From the stress equations (9.7) we then get

$$\widehat{x}x = \frac{k^2}{k^2-1} \operatorname{Re} \left[\psi(z) - \frac{1}{k^2} \psi(z_1) \right], \quad . \quad . \quad . \quad . \quad . \quad (10.8)$$

$$\widehat{x}y = \frac{k^2}{k^2-1} \operatorname{Re} \left[i \left\{ \psi(z) - \frac{1}{k} \psi(z_1) \right\} \right]. \quad . \quad . \quad . \quad . \quad . \quad (10.9)$$

The value of $\widehat{x}y$ on the boundary $y=0$ is given by

$$(\widehat{y}x)_{y=0} = \frac{k}{k+1} \operatorname{Re}[i\psi(z)]. \quad . \quad . \quad . \quad . \quad . \quad (10.10)$$

The value of $\widehat{x}y$ on the boundary being known, we can find the form of the function $\psi(z)$, and hence that of the function $\psi(z_1)$. Substituting the values of these two functions in (10.7), (10.8), and (10.9) we get the required stress components.

XI. Examples.

Case I.

Suppose on the boundary we have

$$\left. \begin{aligned} (\widehat{x}y)_{y=0} &= 0, \\ (\widehat{y}y)_{y=0} &= -p_0 \quad \text{when} \quad -a < x < a, \\ &= 0 \quad \text{for other values of } x. \end{aligned} \right\} \quad . \quad . \quad . \quad . \quad . \quad (11.1)$$

Then we can put

$$(\widehat{yy})_{y=0} = -\frac{p_0}{\pi} \left[\tan^{-1} \frac{y}{x-a} - \tan^{-1} \frac{y}{x+a} \right], \quad \dots \quad (11.2)$$

which gives on substitution in (10.6)

$$f(z) = \frac{p_0(k+1)}{\pi k} i \log \frac{z-a}{z+a} \quad \dots \quad (11.3)$$

Hence we get from (10.3), (10.4), and (10.5)

$$\left. \begin{aligned} \widehat{yy} &= \frac{p_0}{\pi(k-1)} \operatorname{Re} \left[i \left\{ k \log \frac{z_1-a}{z_1+a} - \log \frac{z-a}{z+a} \right\} \right], \\ \widehat{xx} &= \frac{p_0}{\pi(k-1)} \operatorname{Re} \left[i \left\{ \log \frac{z-a}{z+a} - \frac{1}{k} \log \frac{z_1-a}{z_1+a} \right\} \right], \\ \widehat{xy} &= \frac{p_0}{\pi(k-1)} \operatorname{Re} \left[\log \frac{z_1-a}{z_1+a} - \log \frac{z-a}{z+a} \right]. \end{aligned} \right\} \quad \dots \quad (11.4)$$

Case II.

Let us now suppose that on the boundary $y=0$

$$\left. \begin{aligned} \widehat{yy} &= -P \text{ concentrated at the origin, and equal to zero} \\ &\text{elsewhere.} \end{aligned} \right\} \quad \dots \quad (11.5)$$

We have in this case

$$(\widehat{yy})_{y=0} = -\frac{P}{\pi} \frac{y}{x^2+y^2} = -\frac{P}{\pi} \operatorname{Re} \left[\frac{i}{z} \right] \quad \dots \quad (11.6)$$

It is evident from (10.6) that we can put now

$$\left. \begin{aligned} f(z) &= -\frac{P(k+1)}{\pi k} \frac{i}{z^2}, \\ f(z_1) &= -\frac{P(k+1)}{\pi k} \frac{i}{z_1}. \end{aligned} \right\} \quad \dots \quad (11.7)$$

and

Substituting these values in (10.3), (10.4), and (10.5) we get

$$\left. \begin{aligned} \widehat{yy} &= -\frac{P(k+1)}{\pi k^2} \frac{y^3}{r^2 r_1^2}, \\ \widehat{xy} &= -\frac{P(k+1)}{\pi k^2} \frac{xy^2}{r^2 r_1^2}, \\ \widehat{xx} &= -\frac{P(k+1)}{\pi k^2} \frac{x^2 y}{r^2 r_1^2}, \end{aligned} \right\} \quad \dots \quad (11.8)$$

in which

$$r = |z|, \quad \text{and} \quad r_1 = |z_1| \quad \dots \quad (11.9)$$

Case III.

If there be a shear-load only on the boundary such that

$$\left. \begin{aligned} (\widehat{yy})_{y=0} &= 0, \\ (\widehat{xy})_{y=0} &= -q_0 \text{ when } -a < x < a, \\ &= 0 \text{ for other values of } x, \end{aligned} \right\} \quad \dots \quad (11.10)$$

and

we can put

$$\left. \begin{aligned} (\widehat{xy})_{y=0} &= -\frac{q_0}{\pi} \left[\tan^{-1} \frac{y}{x-a} - \tan^{-1} \frac{y}{x+a} \right], \\ &= +\frac{q_0}{\pi} \operatorname{Re} \left[i \log \frac{z-a}{z+a} \right]. \end{aligned} \right\} \quad \dots \quad (11.11)$$

Comparing this result with (10.10), we get

$$\psi(z) = \frac{k+1}{k} \frac{q_0}{\pi} \log \frac{z-a}{z+a} \quad \dots \quad (11.12)$$

The stress components can then be deduced from (10.7), (10.8), and (10.9) in the forms

$$\left. \begin{aligned} \widehat{yy} &= \frac{q_0}{\pi} \frac{k}{k-1} \operatorname{Re} \left[\log \frac{z_1-a}{z_1+a} - \log \frac{z-a}{z+a} \right], \\ \widehat{xx} &= \frac{q_0}{\pi} \frac{k}{k-1} \operatorname{Re} \left[\log \frac{z-a}{z+a} - \frac{1}{k^2} \log \frac{z_1-a}{z_1+a} \right], \\ \widehat{xy} &= \frac{q_0}{\pi} \frac{k}{k-1} \operatorname{Re} \left[i \left\{ \log \frac{z-a}{z+a} - \frac{1}{k} \log \frac{z_1-a}{z_1+a} \right\} \right]. \end{aligned} \right\} \quad \dots \quad (11.13)$$

Case IV.

For a shear load Q concentrated at the origin we have

$$\left. \begin{aligned} (\widehat{xy})_{y=0} &= -\frac{Q}{\pi} \operatorname{Re} \left[\frac{i}{z} \right], \\ &= \frac{k}{k+1} \operatorname{Re} [i\psi(z)] \text{ from (10.10).} \end{aligned} \right\} \quad \dots \quad (11.14)$$

This gives

$$\left. \begin{aligned} \psi(z) &= -\frac{Q}{\pi} \frac{k+1}{kz}, \\ \psi(z_1) &= -\frac{Q}{\pi} \frac{k+1}{kz_1}. \end{aligned} \right\} \quad \dots \quad (11.15)$$

and

Substituting these values in (10.7), (10.8), and (10.9) we get

$$\left. \begin{aligned} \widehat{yy} &= -\frac{Q(k+1)}{\pi k} \frac{xy^2}{r^2 r_1^2}, \\ \widehat{xy} &= -\frac{Q(k+1)}{\pi k} \frac{x^2 y}{r^2 r_1^2}, \\ \widehat{xx} &= -\frac{Q(k+1)}{\pi k} \frac{x^3}{r^2 r_1^2}. \end{aligned} \right\} \dots \dots \dots (11.16)$$

Case V.

If on the boundary we have

$$(\widehat{xy})_{y=0} = 0,$$

$$(\widehat{yy})_{y=0} = -p\sqrt{a^2 - x^2} \text{ when } -a \leq x \leq a,$$

and

$$= 0 \text{ for other values of } x,$$

we can put

$$(\widehat{yy})_{y=0} = -p \operatorname{Re} \sqrt{a^2 - z^2}. \quad \dots \dots \dots (11.18)$$

Comparing this result with (10.6) we get

$$f(z) = -\frac{p(k+1)}{k} \sqrt{a^2 - z^2}. \quad \dots \dots \dots (11.19)$$

Hence from (10.3), (10.4), and (10.5) we obtain

$$\left. \begin{aligned} \widehat{yy} &= -\frac{p}{k-1} \operatorname{Re} [k\sqrt{a^2 - z_1^2} - \sqrt{a^2 - z^2}], \\ \widehat{xy} &= -\frac{p}{k-1} \operatorname{Re} [i\{\sqrt{a^2 - z^2} - \sqrt{a^2 - z_1^2}\}], \\ \widehat{xx} &= -\frac{p}{k-1} \operatorname{Re} [\sqrt{a^2 - z^2} - \frac{1}{k} \sqrt{a^2 - z_1^2}]. \end{aligned} \right\} \dots \dots \dots (11.20)$$

XII. Displacement produced by Loads on the Straight Boundary.

In this section we shall deduce expressions for displacements produced by normal and shear-loads on the straight boundary. From the stress strain relations we have

$$\left. \begin{aligned} \frac{\partial u}{\partial x} &= L \left[\frac{\widehat{xx}}{\sigma_2} - \widehat{yy} \right], \\ \frac{\partial v}{\partial y} &= L \left[\frac{\widehat{yy}}{\sigma_1} - \widehat{xx} \right], \\ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} &= L \left[\frac{1}{\sigma_1} + \frac{1}{\sigma_2} + 2 \right] \widehat{xy}. \end{aligned} \right\} \dots \dots \dots (12.1)$$

(i.) When there is normal load only on the boundary we have from (10.3), (10.4), and (10.5) expressions for the stress components which give

$$\left. \begin{aligned} \frac{\partial u}{\partial x} &= \frac{Lk}{k^2-1} \operatorname{Re} \left[\left(\frac{1}{\sigma_2} + 1 \right) f(z) - k \left(\frac{1}{\sigma_1} + 1 \right) f(z_1) \right], \\ \frac{\partial v}{\partial y} &= \frac{Lk}{k^2-1} \operatorname{Re} \left[\frac{1}{k} \left(\frac{1}{\sigma_2} + 1 \right) f(z_1) - \left(\frac{1}{\sigma_1} + 1 \right) f(z) \right], \\ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} &= \frac{Lk}{k^2-1} \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} + 2 \right) \operatorname{Re} [i\{f(z) - f(z_1)\}]. \end{aligned} \right\} \quad (12.2)$$

Let us now put

$$\int f(z) dz = F_1(z), \quad \text{and} \quad \int f(z_1) dz_1 = F_1(z_1), \quad \dots \quad (12.3)$$

where $F_1(z)$ and $F_1(z_1)$ are analytic functions of z and z_1 respectively.

We can then write the displacement components as

$$\left. \begin{aligned} u &= \frac{Lk}{k^2-1} \operatorname{Re} \left[\left(\frac{1}{\sigma_2} + 1 \right) F_1(z) - k \left(\frac{1}{\sigma_1} + 1 \right) F_1(z_1) \right] - \alpha_1 y + \beta_1, \\ v &= \frac{Lk}{k^2-1} \operatorname{Re} \left[i \left(\frac{1}{\sigma_1} + 1 \right) F_1(z) - i \left(\frac{1}{\sigma_2} + 1 \right) F_1(z_1) \right] + \alpha_1 x + \gamma_1, \end{aligned} \right\} \quad (12.4)$$

where $\alpha_1, \beta_1, \gamma_1$, are arbitrary constants.

(ii.) When there is only shear-load on the boundary, we have on using the results (10.7), (10.8), (10.9)

$$\left. \begin{aligned} \frac{\partial u}{\partial x} &= \frac{Lk^2}{k^2-1} \operatorname{Re} \left[\left(\frac{1}{\sigma_2} + 1 \right) \psi(z) - \left(\frac{1}{\sigma_1} + 1 \right) \psi(z_1) \right], \\ \frac{\partial v}{\partial y} &= \frac{Lk^2}{k^2-1} \operatorname{Re} \left[\frac{1}{k^2} \left(\frac{1}{\sigma_2} + 1 \right) \psi(z_1) - \left(\frac{1}{\sigma_1} + 1 \right) \psi(z) \right], \\ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} &= \frac{Lk^2}{k^2-1} \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} + 2 \right) \operatorname{Re} \left[i \left\{ \psi(z) - \frac{1}{k} \psi(z_1) \right\} \right]. \end{aligned} \right\} \quad (12.5)$$

Putting now

$$\int \psi(z) dz = F_2(z) \quad \text{and} \quad \int \psi(z_1) dz_1 = F_2(z_1), \quad \dots \quad (12.6)$$

where $F_2(z)$ and $F_2(z_1)$ are analytic functions of z and z_1 respectively, we have

$$\left. \begin{aligned} u &= \frac{Lk^2}{k^2-1} \operatorname{Re} \left[\left(\frac{1}{\sigma_2} + 1 \right) F_2(z) - \left(\frac{1}{\sigma_1} + 1 \right) F_2(z_1) \right] - \alpha_2 y + \beta_2, \\ v &= \frac{Lk^2}{k^2-1} \operatorname{Re} \left[i \left\{ \left(\frac{1}{\sigma_1} + 1 \right) F_2(z) - \frac{1}{k} \left(\frac{1}{\sigma_2} + 1 \right) F_2(z_1) \right\} \right] + \alpha_2 x + \gamma_2, \end{aligned} \right\} \quad (12.7)$$

α_2, β_2 , and γ_2 being arbitrary constants.

XIII. *Conclusion.*

In this paper some two-dimensional problems connected with an infinitely large plate having a straight boundary and having different elastic properties along two perpendicular directions are solved. The method employed is the same as followed in previous parts. The results involve a constant $k = \sqrt{\frac{\sigma_1}{\sigma_2}}$, where σ_1 and σ_2 are the Poisson's ratios for contractions along the two perpendicular directions of the axes of x and y respectively. On making k tend to unity we obtain the corresponding results for an isotropic body.

Krishnagar College, Bengal.

LV. *A Matric Theory Development of the Theory of
Symmetric Components.*

By N. S. RISLEY, M.S., and Professor RICHARD S. BURINGTON,
Ph.D., M.A., Department of Mathematics, Case School of
Applied Science, Cleveland, Ohio †.

[Received September 5, 1938.]

1. *Introduction.*

MUCH has been written concerning the theory of symmetric components since the time of Fortescue's original paper ‡. Recently there have appeared a number of papers relating to the development of the theory in matric language. It seems to the authors that a clear and concise presentation of the theory should be made available. In the present paper a synopsis of the theory in matric language is given. The connexion between the recent development by S. Koizumi § is considered.

2. *Linear Correspondence of Vectors.*

Let $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$ be a set of n vectors whose representations with reference to some reference frame \mathbf{F}_1 are the complex numbers

$$\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n, (\mathbf{E}_i = \alpha_i + \beta_i j, j^2 = -1).$$

Similarly, let $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$ be the representations of a second set of n vectors $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$ referred to another reference frame \mathbf{F}_2 .

Let us now refer the set $\{\mathbf{E}\}$ to a new reference frame \mathbf{G}_1 , and the set $\{\mathbf{I}\}$ to a new reference frame \mathbf{G}_2 , and suppose that the relation between the representations $\{\mathbf{E}\}$ in frame \mathbf{F}_1 and the representations $\{\bar{\mathbf{E}}\}$ in frame \mathbf{G}_1 is given by the linear relation

$$\{\mathbf{E}\} = \mathbf{P}\{\bar{\mathbf{E}}\}, \quad d(\mathbf{P}) \neq 0, \quad . \quad . \quad . \quad . \quad . \quad (2.1)$$

while the relation between the representations $\{\mathbf{I}\}$ in frame \mathbf{F}_2 and $\{\bar{\mathbf{I}}\}$ in frame \mathbf{G}_2 is given by

$$\{\mathbf{I}\} = \mathbf{Q}\{\bar{\mathbf{I}}\}, \quad d(\mathbf{Q}) \neq 0, \quad . \quad . \quad . \quad . \quad . \quad (2.2)$$

† Communicated by the Authors.

‡ Fortescue, C. L., "Method of Symmetrical Coordinates applied to the Solution of Polyphase Networks," Trans. A. I. E. E. xxxvii. pt. ii. pp. 1027-1140 (1918).

§ Koizumi, S., "Transformation Formulæ of Impedances and Admittances in the Method of Symmetrical Coordinates," Phil. Mag. ser. 7, xxiv. pp. 195-206 (1937).

This shows that

Theorem 2.1.—The characteristic function $d(A-\lambda I)$ of the correspondence matrix A is an absolute invariant under the similarity (2.6).

It follows immediately that

Theorem 2.2.—The coefficients of λ in the characteristic function of A are absolutely invariant under (2.6).

3. Theory of Symmetrical Components for an n -phase System †.

Consider an n -phase electrical system S having current vectors $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n$, and voltage vectors $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$ in phases 1, 2, \dots , n , respectively. Suppose that these vectors when referred to a reference frame F_1 are represented by the complex numbers I_1, I_2, \dots, I_n and E_1, E_2, \dots, E_n respectively. Then by Kirchhoff's laws an equation for each phase may be written. Suppose these equations are

$$\mathbf{Z}\{\mathbf{I}\}=\{\mathbf{E}\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.1)$$

where $\mathbf{Z}=(Z_{rs})$ is the representation in F_1 of the impedance matrix Z for the system S . Here the elements of \mathbf{Z} are complex numbers. We shall suppose that the system S is linear, so that the impedance elements Z_{ij} in \mathbf{Z} are each constants. We shall further assume that the system S is bilateral, so that \mathbf{Z} is a symmetric matrix, *i. e.*, so that $Z_{ij}=Z_{ji}$.

We shall refer the currents $\{\mathbf{I}\}$ and the voltages $\{\mathbf{E}\}$ to a new reference frame F_2 and suppose that the relation between the representations $\{\bar{\mathbf{E}}\}$ and $\{\bar{\mathbf{I}}\}$ in frame F_2 and the representations $\{\mathbf{E}\}$ and $\{\mathbf{I}\}$ in frame F_1 of the vectors $\{\mathbf{E}\}$ and $\{\mathbf{I}\}$, respectively, are given by

$$\{\mathbf{E}\}=\mathbf{S}\{\bar{\mathbf{E}}\}, \quad \{\mathbf{I}\}=\mathbf{S}\{\bar{\mathbf{I}}\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.2)$$

where $\mathbf{S}\equiv(s_{rt})$, $s_{rt}=a^{-(r-1)(t-1)}$, s_{rt} being the element in the r th row and t th column of \mathbf{S} , and a the primitive n th root of unity. Then (3.1) becomes

$$\bar{\mathbf{Z}}\{\bar{\mathbf{I}}\}=\{\bar{\mathbf{E}}\}, \quad \bar{\mathbf{Z}}=\mathbf{S}^{-1}\mathbf{Z}\mathbf{S}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.3)$$

where $\bar{\mathbf{Z}}$ is the representation of \mathbf{Z} in the new frame F_2 . The rows of \mathbf{S} are referred to by Fortescue as *sequence operators*. It is evident from (3.3) that the equality relationship of *similar matrices* enters the theory of symmetric components.

† Burington, Richard S., "Matrices in Electric Circuit Theory," *Journal of Math. and Physics*, xiv. no. 4, p. 1344 (Dec. 1935).

In accordance with the conventional terminology used by electrical engineers we shall write

$$\{\bar{\mathbf{E}}\} \equiv \begin{Bmatrix} \bar{\mathbf{E}}_1 \\ \bar{\mathbf{E}}_2 \\ \vdots \\ \bar{\mathbf{E}}_n \end{Bmatrix} \equiv \begin{Bmatrix} \mathbf{E}_1^{(0)} \\ \mathbf{E}_1^{(1)} \\ \vdots \\ \mathbf{E}_1^{(n-1)} \end{Bmatrix}, \quad \{\bar{\mathbf{I}}\} \equiv \begin{Bmatrix} \bar{\mathbf{I}}_1 \\ \bar{\mathbf{I}}_2 \\ \vdots \\ \bar{\mathbf{I}}_n \end{Bmatrix} \equiv \begin{Bmatrix} \mathbf{I}_1^{(0)} \\ \mathbf{I}_1^{(1)} \\ \vdots \\ \mathbf{I}_1^{(n-1)} \end{Bmatrix},$$

and shall call the elements $\mathbf{E}_1^{(k)}$ in $\{\bar{\mathbf{E}}\}$ the k th sequence component of voltage in phase 1 ($k=0, 1, \dots, n-1$) and the element $\mathbf{I}_1^{(k)}$ in $\{\bar{\mathbf{I}}\}$ the k th sequence component of current in phase 1.

Since the transformation \mathbf{S} is non-singular, we may solve (3.2) for $\{\bar{\mathbf{E}}\}$ and $\{\bar{\mathbf{I}}\}$, respectively, obtaining

$$\{\bar{\mathbf{E}}\} = \mathbf{S}^{-1}\{\mathbf{E}\}, \quad \{\bar{\mathbf{I}}\} = \mathbf{S}^{-1}\{\mathbf{I}\}. \quad (3.4)$$

These are the sequence components of voltage and current for phase 1. The components for the remaining $n-1$ phases are defined by the relations

$$\{\bar{\mathbf{E}}_r\} = \mathbf{B}\{\mathbf{E}\}, \quad \{\bar{\mathbf{I}}_r\} = \mathbf{B}\{\mathbf{I}\}, \quad (r=1, 2, \dots, n), \quad (3.5)$$

where

$$\{\bar{\mathbf{E}}_r\} \equiv \begin{Bmatrix} \mathbf{E}_r^{(0)} \\ \mathbf{E}_r^{(1)} \\ \vdots \\ \mathbf{E}_r^{(n-1)} \end{Bmatrix}, \quad \{\bar{\mathbf{I}}_r\} \equiv \begin{Bmatrix} \mathbf{I}_r^{(0)} \\ \mathbf{I}_r^{(1)} \\ \vdots \\ \mathbf{I}_r^{(n-1)} \end{Bmatrix}, \quad \mathbf{B} \equiv (b_{ij}), \quad b_{ij} \equiv \delta_{ij} \alpha^{(1-i)(r-1)}.$$

Here δ_{ij} is the Kronecker delta, with $\delta_{ij}=0$ when $i \neq j$, and $\delta_{ij}=1$ when $i=j$. The element $\mathbf{E}_r^{(k)}$ in $\{\bar{\mathbf{E}}_r\}$ is called the k th sequence component of voltage in phase r , and $\mathbf{I}_r^{(k)}$ in $\{\bar{\mathbf{I}}_r\}$, the k th sequence component of current in phase r .

Since the matrix multiplication involved in (3.3) is often rather lengthy a more convenient formula for determining the elements of $\bar{\mathbf{Z}}$ is desired. Koizumi † has devoted a considerable portion of his paper to developing alternate transformation formulæ for obtaining the elements of $\bar{\mathbf{Z}}$, using a system of manipulations on the matrices \mathbf{S} and \mathbf{Z} which is equivalent to a series of elementary transformations on \mathbf{S} and \mathbf{Z} . In this connexion the authors feel that the following formula is simple and practical.

† *Loc. cit.* pp. 197–206.

Since $S^{-1} = \frac{1}{n} S^*$, where S^* is the conjugate matrix to S , we readily see from (3.3) that

$$\bar{Z}_{rt} = \frac{1}{n} P_{r-1} Z Q_{t-1}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.6)$$

where

$$P_k = (1, a^k, a^{2k}, \dots, a^{(n-1)k}), \quad Q_m = \begin{Bmatrix} 1 \\ a^{-m} \\ a^{-2m} \\ . \\ . \\ a^{-(n-1)m} \end{Bmatrix}.$$

A simple method of generating the various sequence components of voltages from the sequence components of voltage in phase 1 is as follows: from $\{\bar{E}\}$ construct the diagonal matrix U , the element in the j th row and k th column of the product $M = US^*$ is the $(j-1)$ th sequence component of voltage for phase k . Thus

$$\begin{aligned} M = US^* &= \begin{bmatrix} \bar{E}_1 & 0 & \dots & 0 \\ 0 & \bar{E}_2 & \dots & 0 \\ . & . & . & . \\ . & . & . & . \\ 0 & \dots & 0 & \bar{E}_n \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & a & a^2 & \dots & a^{(n-1)} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & a^{(n-1)} & \dots & a^{(n-1)(n-1)} \end{bmatrix}, \quad (3.7) \\ &= \begin{bmatrix} E_1^{(0)}, & E_2^{(0)}, & \dots, & E_n^{(0)} \\ E_1^{(1)}, & E_2^{(1)}, & \dots, & E_n^{(1)} \\ . & . & . & . \\ . & . & . & . \\ E_1^{(n-1)}, & E_2^{(n-1)}, & \dots, & E_n^{(n-1)} \end{bmatrix}. \end{aligned}$$

This formula also holds for the sequence components of currents with $\{\bar{E}\}$ replaced by $\{\bar{I}\}$ and $E_k^{(j)}$ by $I_k^{(j)}$.

The system S is said to be *symmetric electrically*, or more briefly *E-symmetric* †, if $Z_{11} = Z_{22} = \dots = Z_{nn}$, and $Z_{sr} = Z_{jk}$ for $r \neq s, j \neq k, r, s, j, k = 1 \dots, n$.

If S is E-symmetric, equation (3.6) yields

$$\left. \begin{aligned} \bar{Z}_{11} &= Z_{11} + (n-1)Z_{12}, \\ \bar{Z}_{rr} &= Z_{11} - Z_{12}, & (r=2, 3, \dots, n), \\ \bar{Z}_{rt} &= 0, & r \neq t, (r, t=1, 2, \dots, n) \end{aligned} \right\}, \quad . \quad . \quad . \quad (3.8)$$

† Risley, N. S., MS, Thesis, Case School of Applied Science, 1938,

so that the impedance matrix representation \bar{Z} in F_2 for the system S reduces to a diagonal matrix. This shows the great simplification that the transformation S produces when the system S is E-symmetric.

An immediate consequence of Theorem 2.2 is that the *trace* of Z is equal to the trace of \bar{Z} ; i. e., that $\sum_{i=1}^n Z_{ii} = \sum_{i=1}^n \bar{Z}_{ii}$. In case S is E-symmetric, we see from (3.8) that the trace is equal to nZ_{11} .

From (3.2) and (3.5), we have the very important result that the sum of the elements of the k th column of M is equal to E_k , that is, the sum of the sequence components of voltage for phase k is equal to the voltage vector E_k . A similar result holds for the components of current. Thus,

$$\sum_{j=0}^{n-1} E_k^{(j)} = E_k, \quad \sum_{j=0}^{n-1} I_k^{(j)} = I^k, \quad (k=1, \dots, n). \quad (3.9)$$

Equation (3.8) shows that when the system S is E-symmetric, sequence currents produce *only* corresponding sequence voltage drops. Thus voltage drops of each sequence may be considered separately, a fact that is not only of great importance from the physical standpoint, but also of distinct advantage from the standpoint of computation by the method of symmetrical components. The sum of the voltage drops of any sequence around any closed circuit must be equal to the e.m.f.'s of that same sequence. Therefore Kirchhoff's second law holds for each sequence of an E-symmetric system S . It is easy to show that each sequence of currents obeys Kirchhoff's first law separately.

It has been shown in the case of an E-symmetric system S that there is no relation between the currents of one sequence and the potentials of a different sequence within the system S . Hence, if zero-sequence currents only flow in the system S , only zero-sequence potential drops will be produced; we shall then designate the system as a *zero-sequence system*. When positive-sequence currents only flow in the system S , only positive-sequence potential drops will be produced; and we shall then designate the system as a *positive-sequence system*. Again, when negative-sequence currents only flow in the system S , only negative-sequence potential drops will be generated; and we shall then designate the system as a *negative-sequence system*.

These sequence systems have been represented by line diagrams, in which the currents are line currents, and the voltages are phase voltages. From electrical considerations it can be shown that the *positive-sequence system* is "structurally identical" with the original system S . The *negative-sequence system* is identical with the positive-sequence system, except that the generated e.m.f.'s are absent. The *zero-sequence system* will in general have fewer branches than the others and will also be free of internal voltages. Since each system is E-symmetric only one phase

need be considered, the interconnexions of the systems depending on the faults or terminal conditions of the system S . The study of unbalanced conditions in a system, caused by unbalanced loads, short circuits or faults on one, two, or more phases, etc., is essentially an engineering problem, so that we will not concern ourselves with it here. The actual construction of "wiring diagrams" for the various sequence systems generated from an arbitrary system of n -phases seems not to have been completely solved.

The method of symmetrical components is evidently applicable to formulæ for admittances as well as impedances. The development parallels the one given above with equation (3.3) replaced by

$$\bar{Y}\{\bar{E}\}=\{\bar{I}\}, \quad \bar{Y}=\mathbf{S}^{-1}\mathbf{Y}\mathbf{S}. \quad . \quad . \quad . \quad . \quad (3.10)$$

We shall now consider in detail the case when $n=3$.

4. 3-phase systems.—In case $n=3$, equations (3.1) are

$$\begin{Bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{Bmatrix} \begin{Bmatrix} I_1 \\ I_2 \\ I_3 \end{Bmatrix} = \begin{Bmatrix} E_1 \\ E_2 \\ E_3 \end{Bmatrix}, \quad . \quad . \quad . \quad (4.1)$$

and equations (3.3) become

$$\begin{Bmatrix} \bar{Z}_{11} & \bar{Z}_{12} & \bar{Z}_{13} \\ \bar{Z}_{21} & \bar{Z}_{22} & \bar{Z}_{23} \\ \bar{Z}_{31} & \bar{Z}_{32} & \bar{Z}_{33} \end{Bmatrix} \begin{Bmatrix} I_1^{(0)} \\ I_1^{(1)} \\ I_1^{(2)} \end{Bmatrix} = \begin{Bmatrix} E_1^{(0)} \\ E_1^{(1)} \\ E_1^{(2)} \end{Bmatrix}, \quad . \quad . \quad . \quad (4.2)$$

where the elements \bar{Z}_{rt} of $\bar{\mathbf{Z}}$ may be obtained from $\bar{\mathbf{Z}}=\mathbf{S}^{-1}\mathbf{Z}\mathbf{S}$, with

$$\mathbf{S} = \begin{Bmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{Bmatrix}; \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.3)$$

or, from (3.6). For example, in case $r=2, t=1, n=3$, we have from (3.6)

$$\begin{aligned} \bar{Z}_{21} &= \frac{1}{3} \mathbf{P}_1 \mathbf{Z} \mathbf{Q}_0 = \frac{1}{3} (1, a, a^2) \begin{Bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{Bmatrix} \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \\ &= \frac{1}{3} (Z_{11} + aZ_{21} + a^2Z_{31}) + \frac{1}{3} (Z_{12} + aZ_{22} + a^2Z_{32}) \\ &\quad + \frac{1}{3} (Z_{13} + aZ_{23} + a^2Z_{33}). \end{aligned}$$

The impedances \bar{Z}_{ik} may be resolved into sequence components in much the same way as the currents and voltages. From the traces of \mathbf{Z} construct the column arrays

$$\{\mathbf{T}_1\} = \begin{Bmatrix} Z_{11} \\ Z_{22} \\ Z_{33} \end{Bmatrix}, \quad \{\mathbf{T}_2\} = \begin{Bmatrix} Z_{12} \\ Z_{23} \\ Z_{31} \end{Bmatrix}. \quad . \quad . \quad . \quad . \quad (4.4)$$

We define the *sequence components of self impedance for phase 1* to be

$$\{Z_{11}^{(p)}\} = \begin{Bmatrix} Z_{11}^{(0)} \\ Z_{11}^{(1)} \\ Z_{11}^{(2)} \end{Bmatrix} = \frac{1}{3} S \{T_1\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.5)$$

and the *sequence components of mutual inductance between phases 1 and 2* to be

$$\{Z_{12}^{(p)}\} = \begin{Bmatrix} Z_{12}^{(0)} \\ Z_{12}^{(1)} \\ Z_{12}^{(2)} \end{Bmatrix} = \frac{1}{3} S \{T_2\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.6)$$

The sequence components of self impedance for phases 2 and 3 are defined by means of the relation

$$\{Z_{rr}^{(p)}\} = B \{Z_{11}^{(p)}\}, \quad (r=2, 3), \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.7)$$

$$\begin{Bmatrix} Z_{22}^{(0)} \\ Z_{22}^{(1)} \\ Z_{22}^{(2)} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a \end{bmatrix} \begin{Bmatrix} Z_{11}^{(0)} \\ Z_{11}^{(1)} \\ Z_{11}^{(2)} \end{Bmatrix}, \quad \begin{Bmatrix} Z_{33}^{(0)} \\ Z_{33}^{(1)} \\ Z_{33}^{(2)} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^2 \end{bmatrix} \begin{Bmatrix} Z_{11}^{(0)} \\ Z_{11}^{(1)} \\ Z_{11}^{(2)} \end{Bmatrix}.$$

The sequence components of mutual impedances for phases 2 and 3 are defined by

$$\{Z_{rk}^{(p)}\} = B \{Z_{12}^{(p)}\}, \quad r \neq k, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.8)$$

that is, by

$$\begin{Bmatrix} Z_{23}^{(0)} \\ Z_{23}^{(1)} \\ Z_{23}^{(2)} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a \end{bmatrix} \begin{Bmatrix} Z_{12}^{(0)} \\ Z_{12}^{(1)} \\ Z_{12}^{(2)} \end{Bmatrix}, \quad \begin{Bmatrix} Z_{31}^{(0)} \\ Z_{31}^{(1)} \\ Z_{31}^{(2)} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^2 \end{bmatrix} \begin{Bmatrix} Z_{12}^{(0)} \\ Z_{12}^{(1)} \\ Z_{12}^{(2)} \end{Bmatrix},$$

$$Z_{jk}^{(p)} = Z_{kj}^{(p)}.$$

The term $Z_{rr}^{(p)}$ is known as the p th sequence component of self inductance in phase r , and $Z_{jk}^{(p)}$ the p th sequence component of mutual inductance between phases j and k .

Using the relations defined above and the fact that $1+a+a^2=0$, we find that $\bar{Z}_{21}=Z_{11}^{(1)}-a^2Z_{12}^{(1)}, \dots$, so that equation (4.2) may be written

$$\begin{bmatrix} Z_{11}^{(0)}+2Z_{12}^{(0)}, & Z_{11}^{(2)}-aZ_{12}^{(2)}, & Z_{11}^{(1)}-a^2Z_{12}^{(1)} \\ Z_{11}^{(1)}-a^2Z_{12}^{(1)}, & Z_{11}^{(0)}-Z_{12}^{(0)}, & Z_{11}^{(2)}+2aZ_{12}^{(2)} \\ Z_{11}^{(2)}-aZ_{12}^{(2)}, & Z_{11}^{(1)}+2a^2Z_{12}^{(1)}, & Z_{11}^{(0)}-Z_{12}^{(0)} \end{bmatrix} \begin{Bmatrix} I_1^{(0)} \\ I_1^{(1)} \\ I_1^{(2)} \end{Bmatrix} = \begin{Bmatrix} E_1^{(0)} \\ E_1^{(1)} \\ E_1^{(2)} \end{Bmatrix}, \quad . \quad (4.9)$$

which gives us the sequence voltage drops for phase 1 only. The voltage drops for phase 2 and phase 3 may be obtained from (3.5); they are

$$\begin{Bmatrix} E_2^{(0)} \\ E_2^{(1)} \\ E_2^{(2)} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a \end{bmatrix} \begin{Bmatrix} E_1^{(0)} \\ E_1^{(1)} \\ E_1^{(2)} \end{Bmatrix}, \quad \begin{Bmatrix} E_3^{(0)} \\ E_3^{(1)} \\ E_3^{(2)} \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^2 \end{bmatrix} \begin{Bmatrix} E_1^{(0)} \\ E_1^{(1)} \\ E_1^{(2)} \end{Bmatrix}. \quad (4.10)$$

If the system S is E -symmetric then (4.9) reduces to

$$\begin{Bmatrix} E_1^{(0)} \\ E_1^{(1)} \\ E_1^{(2)} \end{Bmatrix} = \begin{bmatrix} Z_{11}^{(0)} + 2Z_{12}^{(0)} & 0 & 0 \\ 0 & Z_{11}^{(0)} - Z_{12}^{(0)} & 0 \\ 0 & 0 & Z_{11}^{(0)} - Z_{12}^{(0)} \end{bmatrix} \begin{Bmatrix} I_1^{(0)} \\ I_1^{(1)} \\ I_1^{(2)} \end{Bmatrix}. \quad (4.11)$$

5. Representation of Power in Symmetrical Components.

The power P_s of a system S is given by

$$P_s = (I^*)\{E\} + (I)\{E\}, \quad (5.1)$$

where I^* is the conjugate of I , and $(I^*) = (I^*)S^{*T}$. Here S^{*T} is the transpose of S^* , S^* being the conjugate of S . Equation (5.1) may be written

$$P_s = n \sum_{k=0}^{n-1} I_1^{*(k)} E_1^{(k)} + n \sum_{k=0}^{n-1} I_1^{(k)} E_1^{(n-k) \bmod n}, \quad (5.2)$$

which when $n=3$ reduces to

$$P_s = 3(I_1^{*(0)} E_1^{(0)} + I_1^{*(1)} E_1^{(1)} + I_1^{*(2)} E_1^{(2)}) + 3(I_1^{(0)} E_1^{(0)} + I_1^{(1)} E_1^{(2)} + I_1^{(2)} E_1^{(1)}). \quad (5.3)$$

The real part of $(I^*)E$ is the *average power* P_{av} , and the imaginary part is the *reactive power* Q ; $\sqrt{P_{av}^2 + Q^2}$ is the *vector volt-ampere input*, and the ratio $P_{av}/\sqrt{P_{av}^2 + Q^2}$ is the *power factor*.

6. Sequences in n -phase Systems.

For the general n -phase system a definite relation exists between the number of obtainable sequence systems and the phase order n .

If n is a prime, a system of n vectors $\{E\}$, such as are encountered in an n -phase system, may be resolved into n different sequence systems, one of which consists of n equal vectors and the remaining $(n-1)$ systems consisting of n equi-spaced vectors. If n is composite, that is, when $n = p_1^r p_2^s \dots p_m^q$, where p_1, p_2, \dots, p_m are prime and r, s, \dots, q are positive integers, the system $\{E\}$ may be resolved into a set Γ of n sequence systems of n vectors, and, certain of these sequence systems in Γ will in themselves form the sequence system for a p_i -system (a p_i -phase system), and this is true for $i=1, 2, \dots, m$. Thus in a 6-phase system $\{E\}$ may be resolved into six sequence systems, Γ , of six vectors, and certain of these systems in Γ form the sequence system for a 2-phase system, and certain ones in Γ form the sequence system for a 3-phase system.

LVI. *The Equilibrium Diagram of Iron-nickel Alloys.*

By Prof. E. A. OWEN, M.A., Sc.D., and A. H. SULLY, M.Sc.,
University College of North Wales, Bangor *.

[Received March 30, 1939.]

[Plates VI. & VII.]

1. *Introduction.*

OWING to the importance of iron-nickel alloys in industry a great deal of attention has been devoted to them, and several attempts have been made to establish the equilibrium diagram of the system. One feature that is common to all the diagrams suggested is their comparative simplicity, there being present only three pure phases, namely, the α , γ , and δ , and of these the δ -phase occurs at a temperature in the neighbourhood of 1500°C . This paper will be concerned only with the two phases, α and γ , that is, attention will be directed to temperatures not exceeding about 1000°C . For this region of the equilibrium diagram several arrangements of phase boundaries have been suggested, a summary of which will be found in a report on the alloys of iron and nickel recently published by Desch ⁽¹⁾ in the 'Journal of the Iron and Steel Institute.'

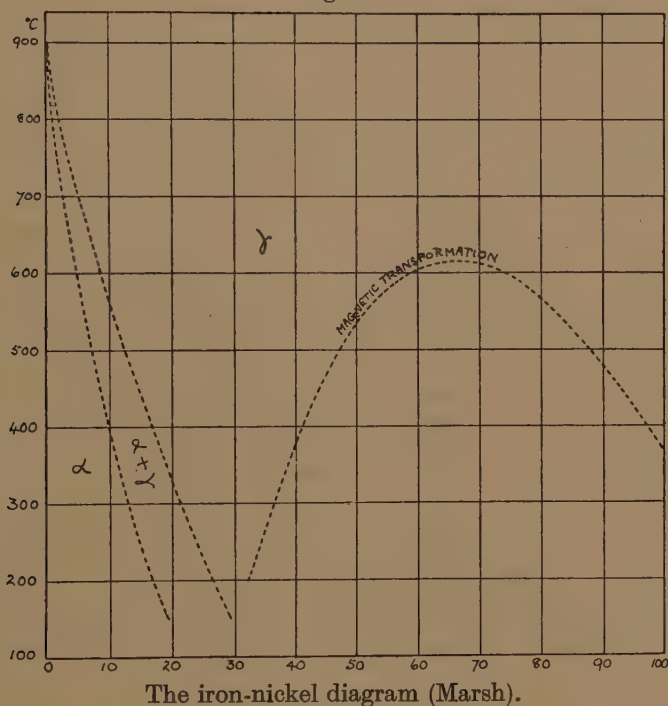
The diagrams of Osmond and Cartaud ⁽²⁾, and of Hanson and Freeman ⁽³⁾ are similar in that they both include a eutectoid change at about 350°C ., and that they incorporate the magnetic-inversion curve to represent a phase change. It is now generally accepted that although it may be desirable to include the magnetic transformation curve in the equilibrium diagram, it does not indicate a true change of phase. In the later diagrams of the system suggested by Honda and Miura ⁽⁴⁾, Merica ⁽⁵⁾, and others, no eutectoid change is indicated, the diagram consisting of the two boundaries $\alpha-(\alpha+\gamma)$ and $(\alpha+\gamma)-\gamma$, and in some cases the region $(\alpha+\gamma)$ exists at different temperatures according as the alloy is heated or cooled, whilst in other cases only one position of the region $(\alpha+\gamma)$ is indicated. In fig. 1 is reproduced the diagram recently suggested by Marsh ⁽⁶⁾ after careful consideration of the available data. The boundaries $\alpha-(\alpha+\gamma)$ and $(\alpha+\gamma)-\gamma$ are shown dotted in this diagram to indicate that the available data do not permit the fixing of the boundaries with certainty. It would appear therefore that, in spite of the great amount of work which has been done on these alloys and the apparent simplicity

* Communicated by the Authors.

of the system, no phase boundaries at the lower temperatures are definitely fixed. The reason for the indecision concerning the location of the phase boundaries is the difficulty experienced in preparing the alloys in a satisfactory state of equilibrium.

The attempt described in this paper to locate the boundaries has been carried out by the aid of X-ray analysis. The alloys are examined in powder form, and there is no doubt that the distortion existing in the particles in their initial condition after filing is an important factor in the problem, as it is found that changes towards the equilibrium conditions

Fig. 1.



occur more rapidly when the material is in powder form than when it is in lump form. But the main factor which inhibits equilibrium at the lower temperatures is the slow rate of the interdiffusion of iron and nickel at these temperatures. This is bound up with another factor which will become clear as the investigation proceeds.

2. Apparatus.

Three X-ray cameras were employed in the investigation : (1) a focusing camera which could be used to obtain precision photographs

of specimens held at any temperature between room-temperature and 550°C. ; (2) a fibre camera used for temperatures up to about 550°C. ; (3) a fibre camera which could be employed with specimens maintained at temperatures ranging up to about 800°C. This camera was especially designed to deal with iron-nickel alloys maintained at high temperatures *in vacuo*, when it became essential to reach temperatures higher than 550°C. , which was about the maximum temperature reached with the fibre camera used hitherto. This camera will be described in detail elsewhere.

3. Heat Treatment.

It is important in an investigation of this nature to pay close attention to the heat treatment of the alloys. Several furnaces were put into commission for the purpose, some for annealing, others for quenching; some were supplied with current direct from the electric mains, others were supplied with current from a battery of large capacity when very steady temperatures were required only for short periods, and still others were supplied with current from the mains, but were controlled thermostatically when constant temperatures were required for long periods.

The material was initially given a definite heat treatment in lump form. Powder specimens were then prepared from the ingots, and these were subjected to further heat treatment. To avoid oxidation the heat treatments, whether the material was in ingot or in powder form, were carried out *in vacuo*, the vacuum being such that, after prolonged treatments at any of the temperatures chosen, the material maintained its metallic lustre.

An extensive survey of the system was made with alloys containing up to about 60 per cent. of nickel and employing widely different heat treatments. Attention will be devoted in this paper to the location of the boundaries in this region of composition by the examination of spectra obtained after these heat treatments, leaving the parameter measurements to be considered in another communication.

4. Effect of Quenching.

The ingots from which the powder samples were prepared, were subjected to two different heat treatments, namely, I., annealed for 12 hours at 1000°C. and quenched into cold water; II., annealed for 19 hours at 1000°C. and slowly cooled to room-temperature over an interval of 39 days. Powder samples from these ingots were subjected to further heat treatments at different temperatures.

(i.) Quenching from 500°C.

The powders were submitted to the following heat treatments: (1) annealed for 48 hours at 500°C. and quenched; (2) annealed for 12 hours

at 900° C., cooled from 900° C. to 500° C. in 3 hours, further annealed for 48 hours at 500° C. and then quenched. The samples were afterwards mounted on a fibre camera and their spectra obtained at room-temperature with cobalt radiation. The phases present are summarized in Table I.

TABLE I.

Nickel content of alloys.		Ingot treatment : either I. or II.	
		Powder treatment.	
Weight (per cent.).	Atoms (per cent.).	(1).	(2).
3.11	2.96	α_1	α_1
6.00	5.73	$\alpha_1 + \gamma$ (trace)	$\alpha_1 + \gamma$ (trace)
8.96	8.56	$\alpha_1 + \gamma$	α_2
12.72	12.17		
14.61	14.01		
16.66	15.98		
21.14	20.32		
24.22	23.32		
28.12	27.12	γ	$\alpha_2 + \gamma$
30.30	29.25	γ	γ
31.67	30.60		
32.20	31.13		
33.59	32.48		
35.10	33.97		
37.17	36.01		
38.42	37.24		
40.66	39.46		
45.50	44.26		
48.00	46.74		

Experiment showed that the results were the same for the two treatments given to the ingots. The behaviour of the alloys in powder form over the range of composition from 8.56 to 23.32 atomic per cent. nickel depended markedly upon the subsequent heat treatments to which they were subjected. The alloys that were annealed for 48 hours at 500° C. and quenched (powder treatment 1) consisted of a mixture of the α - and γ -phases over the range of compositions from 8.56 to 23.32 atomic per

cent. nickel. The α -phase here is designated by α_1 , and its state of equilibrium appears to be comparable with that in the pure α -phase (2.96 atomic per cent. Ni). It is given this designation to distinguish it from the distorted body centred structure designated α_2 found in the same range of alloys which were annealed for 12 hours at 900° C., cooled from 900° C. to 500° C., further annealed for 48 hours at 500° C., and then quenched. Alloy 27.12 is peculiar in that it consists entirely of the γ -phase after treatment (1) and of a mixture of α_2 and γ after treatment (2).

The distorted α_2 structure corresponds to the material which Hanson calls "martensite" owing to its resemblance to martensite in carbon steel, although the structures are not similar—the latter being tetragonal, whilst the former is cubic.

The lines in the X-ray spectra of α_2 are different from those of α_1 in that they are much broader, the (310) doublet of α_2 with cobalt radiation being usually unresolved and diffused. The material is obviously not in a state of equilibrium, and the lattice is considerably distorted.

In those alloys containing less than 27.12 atomic per cent. nickel showing a mixture of two-phases, the γ -phase lines in the spectra were not well defined, thus indicating that this phase also in these alloys was not in equilibrium. In alloys of greater nickel content than 27.12 per cent., the γ -phase on quenching after the two heat treatments mentioned, yielded lines which were very sharply defined, the (222) doublet, which occurs at a large angle of reflexion with cobalt radiation, being well resolved. Thus the γ -phase in these alloys was in a better state of equilibrium than in the other alloys. Reference will be made later to the poor definition of both the α and the γ lines in the mixed region.

The following conclusions may be drawn from the foregoing experiments :—

(1) In alloys annealed at 500° C. and quenched, the γ -phase is completely retained when the nickel content exceeds about 27 atomic per cent. In alloys containing less nickel, both the α - and the γ -phases exist, but they do not appear to be in equilibrium.

(2) Alloys annealed at 900° C. are completely transformed to the γ state. These alloys, when cooled at 500° C. and held at this temperature for some time and then quenched, are found to consist of pure γ when the nickel content exceeds about 29 atomic per cent. In the range of nickel content between 8 and 23 atomic per cent., the γ -phase is completely transformed to a distorted body centred cubic structure (α_2). The production of this distorted α lattice appears to be independent of the length of time (if it exceeds a certain minimum period) that the alloy is maintained at 500° C. before quenching. It is produced after 24 and after 48 hours' annealing at this temperature. Its production depends

on the fact that the temperature is lowered from a higher temperature to 500° C., that is, that the material is in the γ state, before annealing at this temperature, whereas the simultaneous production of the α_1 and γ structures depends on the fact that the alloy is initially in the α condition before being annealed at 500° C. This is an instance of the "irreversibility" associated with alloys in this region of composition.

(3) The boundary of the α -phase at 500° C. lies between 3.0 and 5.73 atomic per cent. nickel, and that of the γ -phase between about 23.5 and 27.5 atomic per cent. nickel. It should, however, be stated that the γ lattice does not appear to be in equilibrium, and possibly the boundary may have to be moved slightly in the direction of greater nickel content.

The positions of the boundaries can be more definitely fixed after observations have been made at different temperatures of annealing.

(ii.) *Quenching from 600° C.*

A similar procedure was adopted with alloys quenched from 600° C. as with alloys quenched from 500° C. The ingots were submitted to the same heat treatments as previously, but the powders were (1) raised to 600° C., maintained at this temperature for 18 hours and then quenched; (2) raised to 900° C. and kept at this temperature for 12 hours, after which they were cooled to 600° C. in 3 hours, then annealed at this temperature for 18 hours and quenched. Two sets of samples covering a range of composition from 2.96 to 50 atomic per cent. were prepared and submitted to the above treatments.

To ensure that equilibrium was attained before the quenching process was performed, other sets of samples were prepared and given twice the above period of annealing at 600° C., namely, 36 hours instead of 18 hours. The results were the same with these samples as with the others.

The results arrived at after examination of the spectra are summarized in Table II. The compositions of the alloys investigated were the same as those shown in Table I., so only the ranges of composition will be included in Table II. except for isolated cases.

Samples from ingots submitted to the two heat treatments gave the same results. It will be observed, however, that the powders of samples with compositions in the range from 8.56 to 23.32 atomic per cent. nickel yield, after heat treatment (1), different results from those obtained when the specimens were annealed at 500° C. The distorted α_2 structure is now obtained after both treatments. Otherwise the results are identical with those obtained after quenching from 500° C.

The conclusions arrived at from these experiments are :—

(1) The α -phase boundary at 600° C. lies between 2.96 and 5.73 atomic per cent. nickel ; it is impossible to say where the γ boundary is situated.

(2) The γ lattice formed at 600° C. is transformed completely on quenching into the "martensitic" α_2 structure in alloys in the range of composition from 8.56 to 23.32 atomic per cent. nickel.

(3) Evidence of irreversibility is found only with alloy 27.12, the γ lattice being retained alone after the first heat treatment, but partially breaking down after the second heat treatment with the formation of the α_2 structure.

TABLE II.

Nickel content of alloys (atoms per cent.).	Powder treatment.	
	(1).	(2).
2.96	α_1	α_1
5.73	$\alpha_1 + \gamma$ (trace)	$\alpha_1 + \gamma$ (trace)
8.56-23.32	α_2	α_2
27.12	γ	$\alpha_2 + \gamma$
29.25-46.74	γ	γ

(iii.) *Quenching from 700° C.*

The samples were raised to a temperature of 700° C., maintained at this temperature for 24 hours and quenched. Another set of samples

TABLE III.

Nickel content of alloys (atoms per cent.).	Phase or phases present after annealing at 700° C. for 24 hours or 7 days and quenching.
2.96-23.32	α_2
27.12	$\alpha_2 + \gamma$
Higher than 29.25	γ

was prepared and maintained at 700° C. for 7 days before quenching. The two sets gave identical results which are summarized in Table III.

Thus the phases present are independent of the time of annealing if this exceeds 24 hours. All the alloys up to a nickel content of 23.32 atomic per cent. have suffered complete conversion to the distorted α_2 form. Alloy 27.12 contained about equal proportions of the α_2 and γ constituents judging from the intensity of the spectral lines. The γ constituent in this alloy appeared to be in a good state of equilibrium,

the (222) doublet being well resolved. Usually when partial conversion to the α_2 form occurs on quenching the retained γ structure yields somewhat broad spectral lines. It is significant also that the 2.96 per cent. alloy gives a distorted α_2 structure whereas, after previous treatments, it yielded the undistorted α_1 structure. Since the α_2 structure is associated with the breakdown of the γ lattice on quenching, it is reasonable to assume that at 700° C. the alloy containing 2.96 atomic per cent. nickel contains the γ constituent. Hence the α -phase boundary at 700° C. corresponds to a composition which is less than 2.96 atomic per cent. nickel.

(iv.) *Quenching from 800° C.*

The results obtained with samples annealed for 24 hours at 800° C. and quenched are included in Table IV.

These results are similar to those obtained after annealing at 700° C. only that now a range of alloys was found in the mixed ($\alpha_2 + \gamma$) region,

TABLE IV.

Nickel content of alloys (atoms per cent.).	Phase or phases present after annealing for 24 hours at 800° C. and quenching.
2.96-23.32	α_2
27.12-30.60	$\alpha_2 + \gamma$
31.13-62.66	γ

so that the pure γ lattice does not appear until the nickel content has reached 31.13 atomic per cent. The γ constituent in the mixed region appears again to be in good equilibrium. The position of the α boundary is again below 2.96 atomic per cent. nickel.

Before considering the above results in detail, further experiments carried out on the α_2 structure and the effect of annealing alloys at lower temperatures and slowly cooling them to room-temperature will be described.

5. *The Distorted Body-centred α_2 Structure.*

(i.) *Effect of Ageing.*

The distorted α_2 structure produced by quenching from high temperatures is obviously a structure which is not in a state of equilibrium and may be subject to considerable internal stresses. It is in an unstable condition, and, in consequence, if alloys in this state were allowed a period of rest some change in the lattice might occur even at room-temperature. Samples of alloys quenched from 600° C., and therefore in the α_2 state,

were examined over a period of about nine months, but no sign of "ageing" was observed, the spectrum lines being equally diffuse at the end and at the beginning of the period of observation, and no γ lines made their appearance. At room-temperature, therefore, the α_2 structure appears to be in a fairly stable condition; a material in a metastable state would behave in this manner.

(ii.) *Effect of Annealing.*

According to the diagram of Merica⁽⁵⁾ alloys in the pure α state at room-temperature do not begin to transform to the γ state on heating until temperatures higher than about 450° C. are attained. If this represents the true state of affairs, alloys in the distorted α_2 state annealed at some temperature below 450° C. should remain in the pure α state and yield the undistorted α form. To investigate this, alloys in the range from 0 to 27.12 atomic per cent. nickel, which were in the distorted α_2 form after quenching from 600° C., were annealed at 400° C. for 21 days, and slowly cooled to room-temperature over a further 12 days. The spectra taken with alloys in the range of composition from 12 to 27 atomic per cent. nickel all showed that the γ lattice was present, the proportion of the γ to the α constituent increasing with increasing nickel content. The distortion of the α lattice was removed by the annealing treatment, the lines being sharp and well resolved at the larger angles of reflexion.

These results seemed to indicate that at 400° C. the mixed phase region extended beyond 27 atomic per cent. nickel, a conclusion which is not in conformity with the diagram of Merica. In view of these observations it was decided to carry the investigation a stage further by studying the effect of slow-cooling the alloys from temperatures not exceeding 500° C. to room-temperature, and of different rates of cooling, on the final constitution of the alloy, thus possibly arriving at the true equilibrium position of the α and γ boundaries.

6. *Effect of Slow Cooling and of Different Rates of Cooling.*

(i.) *Experiments at 500° C.*

Two sets of samples were prepared. The first set was annealed at 500° C. for 66 hours, after which the alloys were air-cooled, reaching air-temperature in 2 or 3 minutes; the second set was annealed at 500° C. for the same period, and slowly cooled to room-temperature over a period of 28 days, that is, cooled at the rate of about 17° C. per day. The spectra obtained with these samples yielded the results collected in Table V.

The alloys included in the ranges in the first column of Table V. were the same as those shown in these ranges in Table I. The γ -phase boundary deduced from the behaviour of the rapidly cooled specimens is situated

at a composition between 23 and 27 atomic per cent. nickel. This is the same result as that already obtained with quenched specimens (Table I.). A comparison of the spectra obtained now with the air-cooled specimens and those obtained with quenched specimens showed that in the former there was a slightly greater proportion of the γ constituent than in the latter, and an improvement in the definition of the lines.

The alloys cooled over 28 days showed an increase in the proportion of the α constituent; this could be detected in alloys up to and including 31.13 atomic per cent. nickel. This result was to be expected, since the slow-cooling process causes the equilibrium resulting in the alloy to correspond to some temperature lower than that from which the cooling was commenced. If an alloy existed in the mixed region in which the atomic binding forces are negligible and the free transference of atoms from one structure to the other takes place at all temperatures, a slow cooling

TABLE V.

Nickel content of alloys (atoms per cent.).	Structure after 66 hours at 500° C. and air-cooled.	Structure after 66 hours at 500° C. and slowly cooled to room-tempera- ture in 28 days.
8.56-23.32	$\alpha + \gamma$	$\alpha + \gamma$
27.12-31.13	γ	$\alpha + \gamma$
32.48	γ	γ

process to room-temperature would yield an alloy in the equilibrium state at the final temperature. Alloys in which the binding forces are appreciable and the transference of atoms sluggish as in the alloys now under investigation need a long period of annealing at any temperature to bring them to equilibrium, and this period increases as the temperature is lowered. In these alloys a slow cooling process such as that described above will result in a state of equilibrium corresponding to some temperature between the initial and the final temperatures. The value of this temperature will depend, amongst other factors, upon the rate of cooling and also upon the mobility of the atoms which depends on the atomic binding forces.

In any iron-nickel alloy in the mixed region the proportion of the α constituent increases as the temperature of annealing is lowered, so that it is to be expected that the slow-cooling process will increase the proportion of the α -phase present in the alloy in the manner which is actually found by experiment. With the rate of cooling adopted the ratio of the amounts of α to γ corresponds to some temperature below

500° C. at which the γ boundary lies between 31 and 32.5 atomic per cent. nickel.

(ii.) *Experiments at 400° C.*

Two sets of samples were prepared, the first set being annealed for 66 hours at 400° C. and air-cooled, the second set annealed for 132 hours at 400° C. and slowly cooled to room-temperature over 28 days. A third set had to be prepared in addition to the above, as it was found that the γ boundary extended beyond the highest composition already included in the previous sets. These were annealed for 132 hours at 400° C. and air-cooled. The spectra yielded the results shown in Table VI.

As with the alloys similarly treated at 500° C., the definition of the lines was better for the slowly cooled than for the air-cooled alloys.

TABLE VI.

Nickel content (atoms per cent.).	Structure after 66 hours at 400° C. and air- cooled.	Structure after 132 hours at 400° C. and cooled over 28 days.
5.73 8.56-32.48	α $\alpha + \gamma$	— $\alpha + \gamma$
	Structure after 132 hours at 400° C. and air- cooled.	
32.29-41.5 44.26-46.74	$\alpha + \gamma$ γ	— —

Unlike the alloys treated at 500° C., the proportions of the α and γ constituents, gauged by the relative intensities of the spectral lines, were the same for the air-cooled and the slowly-cooled alloys. This showed that at temperatures below 400° C. the atomic binding forces must be great and the transformation γ to α very slow.

Alloys of composition between 8.56 and 41.5 atomic per cent. nickel consisted of a mixture of the α and γ constituents, the γ -phase appearing alone when the composition reached 44.26. It is to be noted that the lines in spectra obtained with alloys annealed for 132 hours at 400° C. were sharp and well resolved at the higher angles of reflexion, and were an improvement in this respect over the lines in spectra of alloys annealed for 66 hours at 400° C.

The γ boundary at 400° C. lies between 41.5 and 44.26 atomic per cent. nickel. This result differs greatly from the position of the boundary as

proposed by Hanson and Freeman, who place it at about 19.5 atomic per cent. Osmond and Cartaud show the boundary at about 36 atomic per cent., a result more in agreement with the present work.

The values of the parameters of the α -phase were obtained from spectrum films of alloys containing up to 29.25 atomic per cent. nickel. These are included in Table VII.

The values of the parameters in the duplex region are constant within the limits of experimental error. If the parameter of pure iron is taken to be 2.8607 Å. at 18° C., these values fix the α -phase boundary at 6.5 atomic per cent. nickel. This is in good agreement with the results of Osmond and Cartaud (5.0 per cent.), and with those of Hanson and Freeman (6.0 per cent.).

TABLE VII.

Nickel content (atoms per cent.).	Parameter of α -phase at 18° C.	Nickel content (atoms per cent.).	Parameter of α -phase at 18° C.
2.96	2.8619 Å.	15.98	2.8630 Å.
5.73	2.8629	20.32	2.8632
8.56	2.8629	23.32	2.8631
12.17	2.8628	27.12	2.8631
14.01	2.8630	29.25	2.8631

(iii.) *Experiments at 350° C.*

A number of alloys were annealed at 350° C. for 28 days, others were annealed at this temperature for 89 days, and others for 186 days. All were air-cooled. The spectra obtained with them yielded the results shown in Table VIII.

The original annealing for 28 days was insufficient to bring the alloys into a state of equilibrium, the spectral lines being ill-defined. Those annealed for 89 days showed improvement, but were not as well defined as those in the spectra of alloys annealed for 186 days. But even after this prolonged annealing the lines were still somewhat diffuse, especially those from the γ lattice. Obviously extremely long periods of annealing are required to bring the alloys into equilibrium at these low temperatures. This statement applies more especially to the γ structure, as it was observed after annealing at 400° C. and at 350° C. that the α structure was generally in a better state of equilibrium than the γ structure, the lines in the spectra of the latter being somewhat broad and diffuse when the lines from the α -phase were sharp and well resolved.

A few spectra of alloys annealed at 350° C. are reproduced in fig. 3 (Pl. VI.); the spectrum of alloy 5.73 included in this figure was taken with a camera different from that with which the other spectra were taken. The α lines are clearly seen in all the spectra up to and including that for alloy 44.26. Spectra of other alloys between 44.26 and 50.7 contained α -phase lines, but they were so faint as not to be reproducible. The γ -phase lines in some of the spectra had to be over-exposed in order to bring out the α -phase lines.

The foregoing observations show that the γ boundary at 350° C. lies between about 46.74 and 50.90 atomic per cent. nickel. This result differs greatly from that deduced from Merica's diagram, from which it is concluded that all alloys containing more than 32.5 atomic per cent.

TABLE VIII.

Nickel content (atoms per cent.).	Structure after 28 days at 350° C. and air-cooled.
27.12—37.24	$\alpha + \gamma$
44.26—46.74	Structure after 89 days at 350° C. and air-cooled.
	$\alpha + \gamma$
	Structure after 186 days at 350° C. and air-cooled.
5.73	α
41.50	$\alpha + \gamma$
50.90	γ

nickel are in the pure γ state. Hanson and Freeman show the γ boundary at 350° C. to be at 24.5 atomic per cent. nickel, whereas Osmond and Cartaud give the boundary at about 43 atomic per cent. nickel, neither of which agrees with the present determination.

(iv.) *Experiments at 300° C.*

A set of samples was annealed in a thermostatically controlled furnace for 265 days at 300° C. after which they were air-cooled. An examination of their X-ray spectra yielded the results shown in Table IX.

After this period of annealing the alloys had still not reached their equilibrium state, the γ -phase lines particularly being very diffuse. The observations sufficed, however, to fix the γ boundary approximately ;

the position of the boundary deduced from them would probably be at a composition, the nickel content of which would be too low. Also the α -phase may be present in such small amounts that it would be difficult to determine accurately the limit by this method. This in fact applies to the other determinations also, so that the boundary may have to be moved a little more towards the nickel end of the series on further investigation.

The boundary composition at 300° C. deduced from the above results lies between 50.90 and 57.19 atomic per cent. nickel.

7. Summary of Results with Quenched and Slowly-cooled Specimens.

The foregoing results are shown graphically in fig. 2, which, of course, is not intended to be the equilibrium diagram of this part of the system. The position of the γ boundary cannot be determined from this diagram

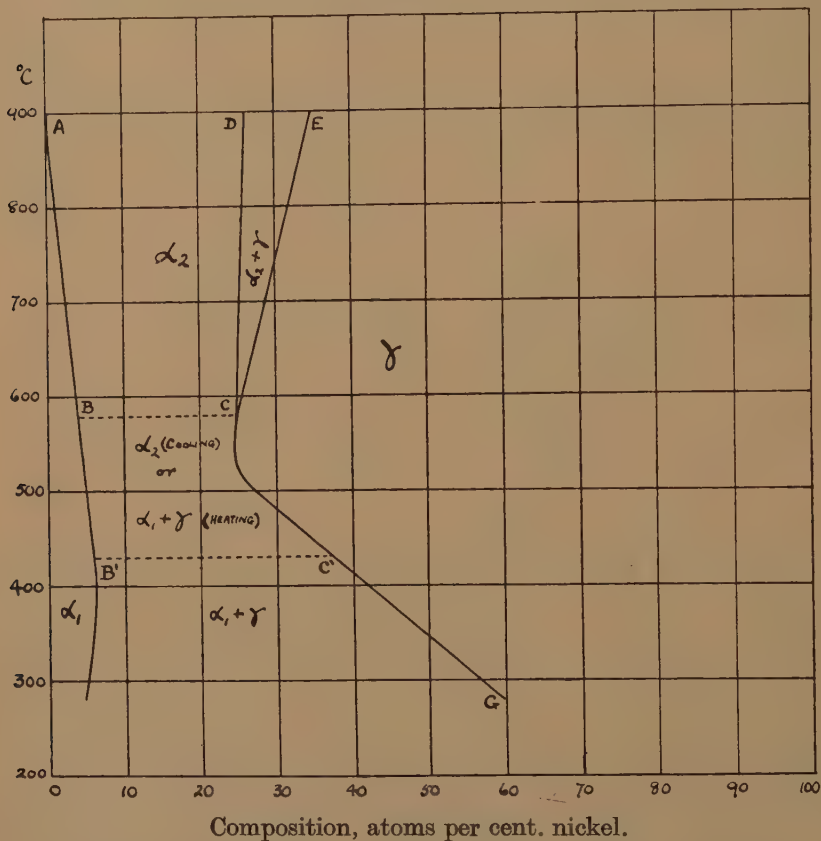
TABLE IX.

Nickel content (atoms per cent.).	Structure after 265 days at 300° C. and air-cooled.
39.46	$\alpha + \gamma$
46.74	$\alpha + \gamma$
50.90	$\alpha + \gamma$
57.19	γ

along its whole length, because above 500° C. the phase existing at these temperatures transforms into a distorted form of the α -phase. But the general direction of the α boundary can be fairly accurately determined, taking into account the position of the boundary at temperatures below 400° C., which has been determined by parameter measurements which will be considered in a later paper. The boundary below about 400° C. shows a tendency to move towards lower nickel content, the nickel content being 6.5 atomic per cent. at 400° C., 5.8 at 350° C., and 4.8 at 300° C. These are approximate values. The values of the composition at various temperatures corresponding to the boundary are shown in Table X., in which are also included the values found by other investigators. There is approximate agreement between the results of the various workers down to about 400° C., but there is no indication of a decrease in the amount of solution of nickel in iron at temperatures below 400° C. in previous investigations. Further investigation is in progress to establish the direction of the α boundary more definitely at temperatures below 400° C.

The results obtained now enable the γ boundary to be fixed for temperatures below 500°C. ; the mean composition at the boundary is given in Table XI., which contains also the corresponding figures recorded by previous workers. The present results agree best with those of Osmond and Cartaud over the range of temperature that is common to the two investigations, but the agreement is far from being satisfactory, and

Fig. 2.



Showing fictitious boundaries obtained with quenched specimens.

whereas Osmond and Cartaud, and Hanson and Freeman found a eutectoid change at about 350°C. , no such change is now observed. The results of Hanson and Freeman, and of Merica differ widely from the results recorded here, the amount of solution of nickel in iron at different temperatures being now found to be more than double that recorded by these investigators.

Jette and Foote ⁽⁷⁾ determined by X-ray analysis a few points on the α and the γ boundaries. These are included in Tables X. and XI. The

TABLE X.
Position of the α boundary.

Temperature (°C.).	Nickel content (atoms per cent.) at boundary.				
	Osmond & Cartaud.	Hanson & Freeman.	Merica.	Jette & Foote.	Present work.
800	0.0	2	1	—	1.0
700	2.0	2	2	—	2.5
672	—	—	—	3.94	2.7
600	3	3	4	—	3.5
557	—	—	—	4.22	4.0
500	4	4	6	—	5.0
496	—	—	—	6.07	5.0
456	—	—	—	5.91	5.5
400	5	6	8	—	6.5
350	—	—	—	—	5.8
300	—	—	—	—	4.8

TABLE XI.
Position of the γ boundary below 500° C.

Temperature (°C.).	Nickel content (atoms per cent.) at boundary.				
	Osmond & Cartaud.	Hanson & Freeman.	Merica (on cooling).	Jette & Foote.	Present work (± 0.5).
500	26.0	13.1	10.0	—	27.0
456	—	—	—	33.36	34.0
450	31.2	16.1	11.7	—	34.5
400	36.8	19.6	13.6	—	42.0
350	43.4	24.5	15.3	—	49.0
300	—	—	16.9	—	56.5

point on the γ boundary at 456° C. agrees closely with that found at present. On the whole the points on the α boundary also agree with the present results, the greatest difference being 1.2 atomic per cent. nickel at 672° C.

It is interesting to note that a second point determined by Jette and Foote on the γ boundary at 585°C . corresponds with 26.20 atomic per cent. nickel, which agrees closely with the boundary composition (25.5 atomic per cent. nickel) at this temperature shown in fig. 2, which does not represent equilibrium conditions.

The diagram in fig. 2 shows the behaviour of alloys which have been annealed at temperatures above 500°C . and then quenched. The γ lattice in alloys in the region ABCD is converted completely on quenching to the distorted α_2 structure, but in the region DCE the γ lattice is only partially converted to the distorted α_2 form. The γ lattice is completely retained on quenching only in alloys with nickel contents greater than those defined by the line GCE.

The line BC, which has only been approximately fixed, represents the temperature above which the γ lattice is completely destroyed on quenching and below which some of the γ structure is retained. For alloys raised from room-temperature to the temperature from which they are quenched, BC lies, as shown in the diagram, between 500°C . and 600°C ., quenching from the former temperature yielding the mixed ($\alpha+\gamma$) structure and from the latter temperature the distorted α_2 structure.

If, however, the temperature is first raised to a high value (say 900°C .), and then lowered to the temperature from which the quenching takes place, the line BC moves to some position such as B'C' between 500°C . and 400°C ., since now the distorted α_2 structure is produced on quenching from 500°C ., whereas a mixture of the α - and γ -phases results on quenching from 400°C . The region BB'C'C is peculiar therefore in that it contains alloys which yield on quenching different results depending upon their previous heat treatment. The determination of more definite positions for BC and B'C' is still to be carried out. The region BB'C'C may be called the region of irreversibility.

In fig. 4 (Pl. VII.) are reproduced some of the spectra obtained with an alloy containing 16.0 atomic per cent. nickel. These spectra show the existence of the α_2 structure in alloys quenched from temperatures between 600°C . and 800°C . There seems to be an improvement in the definition of the lines as the temperature of annealing is raised. The spectra of the alloy when annealed at 500°C . show the effect of slow cooling on the definition of the lines. The spectra give an indication also of the state of equilibrium of the alloy after annealing at such low temperatures as 400°C . and 350°C .

The fact that the γ lattice is transformed completely to the α state in alloys quenched from temperatures above 500°C . made it impossible to determine the γ boundary at these temperatures. The only method by which this could be accomplished was by taking X-ray spectra of the alloys whilst they were maintained at the required temperature. This

necessitated the construction of a camera in which the alloys could be maintained at elevated temperatures for prolonged periods extending over the time required for annealing and for taking the photograph.

8. Observations at High Temperatures.

The powder samples were taken from the ingots which had been submitted to the heat treatments already mentioned. They were fixed on the fibre with a very dilute solution of adhesive. The fibre was placed in the holder of a camera which was mounted in an evacuated enclosure, the vacuum being maintained at about 0.001 mm. pressure throughout the experiment.

TABLE XII.

The γ boundary of the Nickel-iron System.

Temperature °C.	Nickel content (atoms per cent.) at the γ boundary.
800	4.5
700	9
600	14
550	18.5
500	26.5
450	34
400	41.5
350	49
300	56.5

The specimen under examination was first of all heated to the required temperature in the enclosure, and maintained at this temperature for 24 hours before commencing the exposure, which lasted between 1 and 2 hours. Temperatures up to about 800° C. were reached with this new design of camera. The method consisted in taking spectra of a series of alloys and in finding at what composition the structure changed from the pure γ -phase to the mixed ($\alpha + \gamma$) phase.

It was soon found that alloys which previously gave the distorted α_2 structure on quenching now consisted of the pure γ -phase at high temperature. The position of the γ boundary between 500° C. and 800° C. was satisfactorily determined in this manner. The compositions corresponding to the γ boundary at various temperatures are given in Table XII.

*9. The Equilibrium Diagram of the Iron-nickel System.**(i.) General Considerations.*

The equilibrium diagram of the system drawn in fig. 5 represents the results so far obtained in this investigation. The α boundary is fairly definitely established, but the γ boundary may have to be moved a little further towards the nickel end of the diagram. The magnetic transformation boundary is also included, but is drawn dotted, as it does not form part of the phase diagram.

The diagram now put forward differs from previous diagrams in at least two respects, namely, (1) the α boundary shows a maximum of solubility of nickel in iron at about 400° C., the solubility gradually decreasing at lower temperatures; (2) the slope of the γ boundary below 500° C. is smaller than in previous diagrams, in other words, the increase in the solubility of iron in nickel with rise of temperature is greater than that found previously.

The change in solubility of nickel in iron between 500° C. and 300° C. is only about 2 atomic per cent., whereas the change in the solubility of iron in nickel over the same range of temperature is not less than about 30 atomic per cent.

A result which emerges immediately from the examination of the X-ray spectra of alloys in the duplex region ($\alpha + \gamma$) is that the α -phase can be more readily obtained in a state of equilibrium than the γ -phase. This can be explained in the light of the equilibrium diagram now suggested for this part of the system. The saturation value of the solution of nickel in iron changes so little with temperature over the range 500° C. to 300° C., that any slight fluctuation in temperature would have no appreciable effect on the crystal lattice of the α -phase, but this fluctuation would have marked effect on the crystal lattice of the γ -phase. Or if the alloy is subjected to slow cooling, say from 500° C. to room-temperature, the saturation composition of the α -phase changes little and approximate equilibrium conditions are readily established; but for the γ -phase the saturation composition changes rapidly, the iron constituent being removed from solution more rapidly than it can be absorbed or than the γ lattice can settle down in its equilibrium state. Exceedingly slow cooling under well-controlled conditions is necessary to produce a γ lattice that is completely in equilibrium. At these low temperatures the inter-diffusion of iron and nickel is so slow that a very long period of cooling would be necessary.

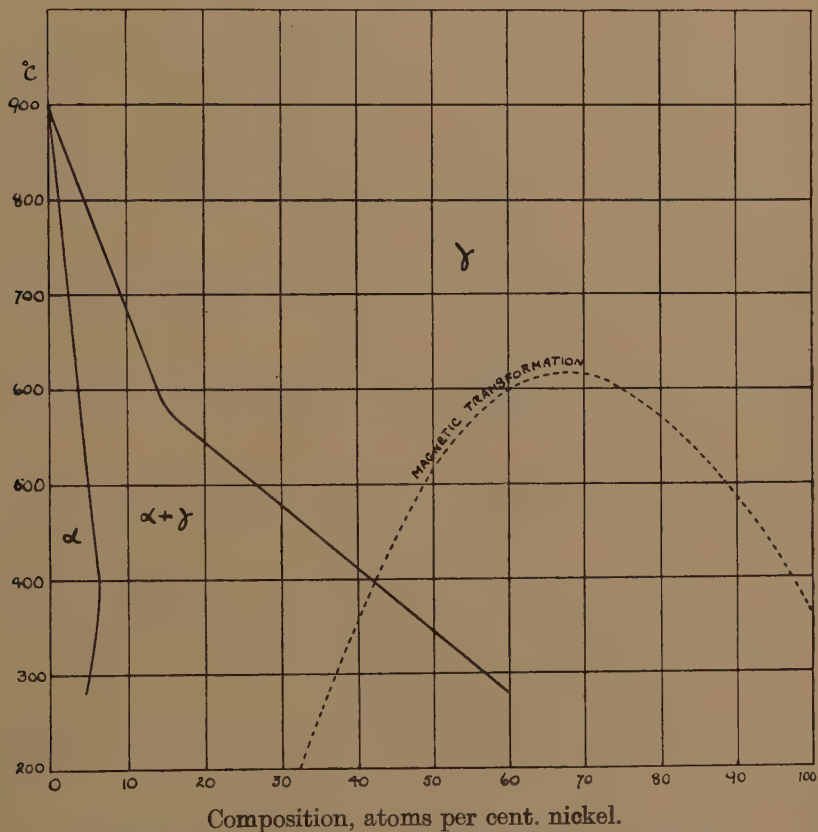
It is an interesting experimental observation in this connexion that in alloys the compositions of which fall to the right of the line ECG in fig. 2, and which are quenched, the γ -phase is found to be in a good state of equilibrium; this is not so for the γ -phase in quenched alloys belonging

to the $(\alpha+\gamma)$ region. In other words, the definition of the γ -phase lines shows a marked improvement when the boundary ECG is crossed from left to right.

(ii.) $\alpha \rightarrow \gamma$ and $\gamma \rightarrow \alpha$ Transformations.

It has long been known that the $\alpha \rightarrow \gamma$ transformation takes place much more readily than the $\gamma \rightarrow \alpha$ transformation, and this has been difficult

Fig. 5.



Equilibrium diagram of iron-nickel system below 1000° C.

to explain in the light of previous equilibrium diagrams. The present results suggest an explanation of this effect. The α region is limited within the region 0 to about 6 atomic per cent. nickel. As the temperature of any alloy in the pure α -range is raised, the solubility of nickel in iron gradually decreases, but is almost stationary between 300° C. and 500° C. The nickel that gradually comes out of solution as the temperature is

raised above 500°C . is readily accommodated in the γ -phase at these higher temperatures until the transformation is complete.

In the other direction the conditions are very different. Starting with the pure γ -phase in the region between, say, 6 and 25 atomic per cent. nickel, the temperature will have to be lowered at a much slower rate, as already indicated, to maintain even approximate equilibrium conditions. For this reason the transformation of γ to α is found to be much more difficult than the reverse transformation.

Somewhat similar reasons may be put forward to explain the existence of the distorted body-centred α_2 structure in alloys quenched from temperatures above about 500°C . The high temperature observations show that the γ -phase is the equilibrium state of alloys at these temperatures. On quenching, the γ -phase breaks up forming a body-centred structure, the nickel becoming randomly distributed in the iron lattice which forms when the iron is removed from solution owing to the rapid decrease in temperature. The body-centred lattice thus formed contains excess nickel and is distorted; it may be regarded as a supersaturated solid solution of nickel in iron. It breaks up into a mixture of α - and γ -phases in equilibrium when annealed again at temperatures below 500°C . The lack of definition of the lines in the spectrum of the α_2 structure may be due partly to broadening on account of the submicroscopic size of the crystals and partly to distortion of the lattice.

(iii.) *Interdiffusion below 500°C .*

Marsh ⁽⁶⁾ states that the general opinion is that iron-nickel alloys in the "irreversible" region "freeze in" at about 500°C . on cooling, that is "at all temperatures below 500°C . atomic diffusion either ceases or becomes negligibly small, consequently equilibrium is either impossible or extremely difficult to attain." In the work described here, which was carried out with powder specimens, interdiffusion of iron and nickel atoms does take place when specimens, which on quenching from high temperatures formed the distorted α lattice, are annealed at temperatures below 500°C . This migration of atoms is no doubt due in some measure to the fact that the lattice is distorted, and therefore is subjected to internal stresses.

In the course of the work it has been observed that, in general, filings, owing to their distorted condition when prepared, show changes in lattice formation much more readily than a lump of the material. Whether the distortion is produced mechanically or by quenching, subsequent heat treatment has much more effect on the lattice than if no distortion were present.

Hence a statement to the effect that atomic diffusion either ceases or becomes negligibly small at temperatures below 500°C . needs qualification if there is distortion present.

(iv.) The Magnetic-inversion Curve.

The magnetic-inversion curve is shown dotted in fig. 5. It cuts the γ -phase boundary at about 42 atomic per cent. nickel, and runs into the mixed ($\alpha + \gamma$) region. The magnetic properties of the alloys prepared for the present investigation have not been examined. Attention may, however, be drawn to the measurements made by various investigators on the saturation value of magnetization of γ -phase alloys at ordinary temperatures. A curve showing the relation between the saturation magnetization and the nickel content will be found in Marsh's book⁽⁶⁾, in which is incorporated the results of several investigators. The curve shows a steady increase in the saturation value of the magnetization as the nickel content decreases, and, as drawn, it reaches a maximum of about 16,000 gauss at about 46 to 47 atomic per cent. nickel, but on further decrease of nickel content, the saturation magnetization diminishes rapidly, reaching a value of about 2000 gauss at about 29 atomic per cent. nickel. If the results are examined individually, the maximum is found to occur approximately at 42 atomic per cent. nickel by Peschard⁽⁸⁾, at 41 per cent. by Masumato⁽⁹⁾, and at 43 per cent. by Lichtenberger⁽¹⁰⁾. These values are close to the nickel content at the intersection of the magnetic-inversion curve and the γ boundary, namely, 42 atomic per cent. nickel. If the mean position of the maximum is taken at this composition, it may then be stated that if, on lowering the temperature, the magnetic-inversion occurs before the γ -phase boundary is crossed, the saturation value of the magnetization gradually increases with decrease in nickel content of the alloy, and reaches a maximum for the composition corresponding to the point at which the two boundaries cross. On the other hand, if on lowering the temperature, the γ -phase boundary is crossed before the magnetic-inversion takes place, a very rapid decrease in the saturation value of the magnetization occurs as the nickel content is lowered below that corresponding to the maximum.

10. Acknowledgments.

The authors take this opportunity to record their indebtedness to the Research and Development Department of the Mond Nickel Company for their assistance in supplying material for the investigation, and particularly to Dr. L. B. Pfeil for his interest in the work and for information concerning the alloys.

11. Summary.

The paper contains an account of the determination by X-ray methods of the positions of the α and γ boundaries in the equilibrium diagram of the iron-nickel system of alloys. The alloys were given careful heat treatments, and the effects of these treatments on the structures of the

alloys were studied. Spectra were taken with material maintained *in vacuo* at high temperatures reaching up to about 800° C., a special design of high temperature camera being used for the purpose.

An attempt is made with the aid of the proposed new equilibrium diagram to explain certain experimental facts which were difficult to account for on the basis of previous diagrams.

References.

- (1) C. H. Desch, Iron and Steel Inst., First Report of the Alloy Steels Research Committee, Sect. V. (1936).
- (2) F. Osmond and G. Cartaud, *Revue de Metallurgie*, i. p. 69 (1904).
- (3) D. Hanson and J. R. Freeman, Journ. Iron and Steel Inst. i. p. 301 (1923).
- (4) K. Honda and S. Miura, Tohoku Univ., Sci. & Technol. Reports, xvi. p. 745 (1927).
- (5) P. D. Merica, 'National Metals Handbook,' p. 339 (1933 ed.).
- (6) J. S. Marsh, 'Alloys of Iron and Nickel,' i. p. 54 (1938 ed.).
- (7) Jette and Foote, Trans. Am. Inst. Min. Met. Eng. cxx. p. 259 (1936).
- (8) M. Peschard, *C. R. Acad. Sci.* cl. xxxi. p. 854 (1925).
- (9) H. Masumato, Sci. Report, Sendai, xviii. p. 195 (1929).
- (10) F. Lichtenberger, *Ann. d. Physik*, xv. p. 45 (1932).

LVII. *The Disintegration of Erythrocytes and Denaturation of Hemoglobin by High Pressure.*

By R. B. DOW, A.M., Ph.D., and J. E. MATTHEWS, Jr., A.B.,
Departments of Physics and Agricultural and Biological Chemistry,
The Pennsylvania State College *.

[Received November 14, 1938.]

[Plates VIII.—X.]

No data appear in the literature on the general effects of high hydrostatic pressure on mammalian blood, although Bridgman and Conant † have reported that Dr. A. R. Davis observed the coagulation of blood at high pressure several years prior to their experiments on the apparent denaturation of carboxyhemoglobin at pressures of 9000 atmospheres.

We have observed recently several interesting pressure effects in bovine blood which have not been reported hitherto. These experiments were performed in the high pressure laboratory of the Department of Physics with apparatus similar to that of Bridgman and Conant. The samples of blood, prepared as described below, were subjected to constant pressures that ranged from 3500 to 13,000 atmospheres at room-temperature which varied from 27° to 29° C. Two series of experiments are described and the results briefly summarized.

I.

Initial tests were made with sterile oxalated, bovine blood. Five to seven c.c. of blood were introduced into a glass tube by evacuating the tube under blood within a bell-jar and then admitting air. The tube was then inverted in mercury in a steel cylinder and the assembly placed in the testing chamber of the high pressure apparatus.

A typical experiment was one run at 3500 atmospheres for 3 hours with the following results:—erythrocyte count decreased from 7,300,000 to 800,000, and leucocytes decreased from 6700 to 800 per cu. mm., both approximately in the same proportion; both types of cell appeared distorted in a microscopic field of 1000 x ; NaCl crystals, and possibly crystals of other salts, were clearly visible; oxygen binding capacity decreased by 41 per cent. as measured by the Van Slyke method. Pressure treatment of a sample for 1 hour at 3500 atmospheres gave results roughly

* Communicated by the Authors.

† P. W. Bridgman and J. B. Conant, Proc. N. A. S. xv. p. 680 (1929).

comparable with those of the 3 hour exposure. The oxygen binding capacity in this case decreased by 16 per cent.

A 6 hour exposure at 3500 atmospheres, as well as a $3\frac{1}{2}$ hour exposure at 13,000 atmospheres, coagulated the samples of blood into reddish-brown rubbery masses. Microscopic examination of the precipitate revealed no erythrocytes, very few distorted leucocytes, no crystallization, and a jelly-like material that was without apparent structure but containing dispersed nuclei from leucocytes.

II.

The second series of tests were made on blood-cells separated from freshly drawn oxalated bovine blood by means of a centrifuge, and with aqueous solutions of hemoglobin in order to corroborate the results of (I.) under more simple experimental conditions. The aqueous solutions contained approximately 50 per cent. normal concentration of hemoglobin. The samples in these tests were placed in paraffin-lined steel tubes without filling by evacuation. Paraffin-oil was used to separate the pressure-transmitting liquid from the samples. The experiments were conducted at constant pressures of 5000 atmospheres for intervals ranging from $\frac{1}{2}$ to 2 hours.

The blood-cells in each case were entirely coagulated, giving a material similar to that observed previously (I.) when coagulation was obtained at higher pressures and for longer times. Apparently the proteins of the red cells, due either to their higher concentration or to their nature, were more easily coagulated than the proteins of the plasma. No crystallization occurred in this series of experiments.

The aqueous solutions of hemoglobin were treated with the same pressures for periods of 2 and 4 hours respectively. The coagulation, while not complete, was similar in structure to that observed for bovine blood. It appeared similar to that obtained by treating control solutions with ethyl alcohol and with diethyl ether. Figs. 1, 2 (Pl. VIII.) and fig. 3 (Pl. IX.) are photomicrographs ($1000\times$) of blood-cells exposed at 5000 atmospheres for $\frac{1}{2}$ hour, aqueous solution of hemoglobin exposed at 5000 atmospheres for 2 hours, and control hemoglobin solution treated with alcohol, respectively. Bridgman and Conant believe that carboxyhemoglobin is "denatured" by pressure treatment similar to the denaturation obtained by the action of alcohol on carboxyhemoglobin. The result of our experiments lead us to much the same conclusion for the effect of pressure on hemoglobin.

To observe the effect of pressure on the erythrocytes more clearly a small sample of blood-cells was suspended in Hayem's solution and then subjected to pressure treatment of 5000 atmospheres for 1 hour. Fig. 4 (Pl. IX.) and fig. 5 (Pl. X.) are photomicrographs ($1000\times$) of the

cells before and after pressure treatment respectively, showing the drastic effects of pressure treatment.

In summarizing the results of the two series of tests it is interesting to observe that crystallization of salt occurred when the coagulation of protein was incomplete. This would indicate that salts were being removed from solution by reacting with other substances present under pressure, and agree with the views of other investigators * who believe that in a denaturation process the denaturation may inhibit crystallization.

The question arises whether the observed disintegration of blood-cells is due to the mechanical effects of pressure or chemical effects under pressure. This was partly explained by results obtained when an unexpected leak developed in the pressure apparatus during one of the tests. After an exposure of about 1 minute at 2500 atmospheres a leak in pressure connexions was noticed, and in an attempt to maintain the pressure by rapid pumping considerable fluctuation in pressure necessarily occurred. Microscopic examination of the blood then showed numerous salt crystals in the field but little change in the appearance of the cells. The nuclei of the leucocytes suffered elongated distortion. This test indicated to us that the effects of pressure change were not primarily mechanical ones. While the cells are clearly distorted and might be ruptured by quick release of pressure it is not likely that the pressures that we used could be responsible for the observed disintegration of cells.

As mentioned previously, one of the effects of pressure seemed to be comparable to the addition of an OH radical to the hemoglobin. This effect may involve the breaking up of polypeptide linkages $\text{>N}-(\text{OH})\text{C}<$ into new independent groups, >NH , $\text{OC}<$ possibly accompanied by hydration. This mechanism of denaturation proposed by Wrinch † seems extremely suggestive, but we believe that pressure aids the disintegration rather than inhibits it as Wrinch predicts. The reaction under pressure appeared to be a reduction. The disintegration of erythrocytes resulted from this reaction rather than from rupture of cell structure. The decrease of O_2 binding capacity content appears to be governed by a first-order reaction, similar to the observations of Bridgman and Conant on CO.

This research was made possible by the co-operation of several of the senior author's colleagues. Dr. J. F. Shigley generously supplied samples of bovine blood, Dr. M. L. Willard loaned the camera for the photomicrographs, and Dr. A. K. Anderson not only placed his equipment at our disposal but also assisted in criticizing the manuscript.

* D. M. Wrinch, *Phil. Mag.* xxv. p. 705 (1938).

† *Loc. cit.*

LVIII. *Equivalence of two Piezoelectric Oscillating Quartz Crystals of Symmetrical Outlines with Respect to a Plane Perpendicular to an Electrical Axis.*

By ISSAC KOGA, Doctor of Engineering
(Tokyo University of Engineering)*.

[Received December 5, 1938.]

1. *Theorem.*

IN cutting out piezoelectric oscillating plates from quartz crystal in various orientations and outlines, the following new theorem seems to serve as a criterion to determine whether two plates are theoretically equivalent or not :

Two piezoelectric oscillating crystals, symmetrical in their outlines and in their relationship to the electrodes referring to a plane perpendicular to an axis of symmetry excepting a trigonal axis, are equivalent ;

and as a special case :

Two piezoelectric oscillating quartz crystals, symmetrical in their outlines and in their relationship to the electrodes referring to a plane perpendicular to an electrical axis, are equivalent.

2. *Proof.*

A brief explanation will now be given, citing one or two concrete examples.

As is well known, a crystal has always a centre of symmetry in its elastic behaviour, and when there exist both a centre of symmetry and an axis of symmetry other than a trigonal axis, it has always a plane of symmetry perpendicular to the axis of symmetry, so that as far as the elastic behaviour of a crystal is concerned there exists always a plane of symmetry perpendicular to an axis of symmetry other than a trigonal axis.

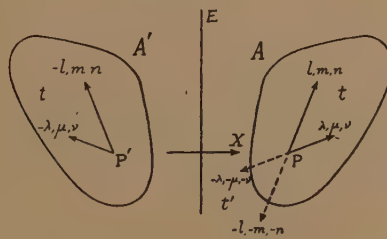
In the same manner, it is easily seen that the same theorem can be applied to the frequencies and modes of vibrations of two elastic bodies which are symmetrical with respect to a plane perpendicular to an axis of symmetry other than a trigonal axis.

* Communicated by the Author.

Now take into account the piezoelectric effect in addition to the above consideration. Suppose there are two piezoelectric crystals A and A' , as shown in fig. 1, which are symmetrical to each other with respect to a plane mirror E perpendicular to an axis of symmetry other than a trigonal axis, taken as X axis of rectangular co-ordinate axes; and suppose an arbitrary point P in a body A is displaced to an orientation of direction cosines λ, μ, ν at any instant t . Then the point P is displaced in the opposite direction $-\lambda, -\mu, -\nu$, and at the same time strains, stresses, electric polarization due to piezoelectric effect, etc., are all reversed after half a period of vibration, say at time t' . Therefore, if the body A is first electrically polarized at time t in the direction of l, m, n at P , the polarization at the same point P will be reversed to the direction of $-l, -m, -n$ at time t' .

On the other hand, as the axis X is an axis of symmetry other than a trigonal axis, the piezoelectric effect remains unchanged if all are

Fig. 1.



rotated by 180° about the axis X , while the displacements, strains, and stresses at point P at time t' thus rotated are nothing but the images on the mirror E of those which occurred at point P at time t in A . Consequently, it can be concluded that the piezoelectric polarization at P' due to elastic vibration of A' , which is the exact image of elastic vibration of A , is equal and opposite in sense to the image on E of polarization at P in A .

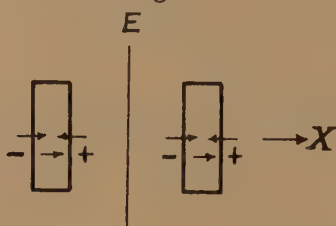
Therefore, so long as a piezoelectric crystal is used as an element of alternating-current circuit individually, it should be completely equivalent to another one which is symmetrical in the outline and in the relationship to its electrodes referring to a plane perpendicular to an axis of symmetry other than a trigonal axis, because the relationship between the mode of vibration and the direction and magnitude of piezoelectric polarization in the latter is exactly equal to that in the former. Thus the theorem is proved.

3. Corollary.

Since a piezoelectric quartz may thus be regarded as though it has planes of symmetry perpendicular to the digonal (electrical) axes, as

described above, so long as an oscillating plate is treated individually it may be considered as a corollary of this theorem that there would be no more distinction between right-handed and left-handed quartz. Accordingly, any piezoelectric oscillating quartz can always be uniquely determined by merely having its relationship to r -face (positive or direct rhombohedron of the first order), r' -face (negative or inverse rhombohedron of the first order), and m -face (hexagonal prism of the first order) clearly identified, and it is not necessary to know whether the quartz is right-handed or left-handed.

Fig. 2.



4. Example.

For example, take a case of two circular quartz plates, shown in fig. 2. It is quite evident that these two plates, if cut exactly the same in size and

Fig. 3.

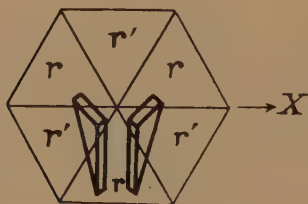
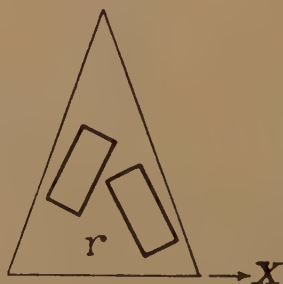


Fig. 4



orientation (perpendicular to an electrical axis X), are identical. However, this may be considered as one of the cases to which the foregoing theorem can be applied. That is, when these two plates, symmetrical in outlines with respect to a plane E perpendicular to an electrical axis X , are vibrating, for instance, in the longitudinal mode in the X direction and in the same phase, the electrical polarization will be produced in the same orientation. This means that the mechanical vibration of one plate coincides exactly with the image on the mirror E of the other, and the electrical polarization is equal and opposite in sense to the image.

Therefore, so long as the two plates are used individually, they are exactly equivalent.

In both figs. 3 and 4 are shown two oscillating plates having symmetrical outlines with respect to a plane perpendicular to an electrical axis X . So the two oscillating plates are equivalent when their electrodes are also placed symmetrically with respect to the same plane, such as, for instance, when the two electrodes are placed to face the principal surfaces.

5. *Summary.*

Two piezoelectric oscillating quartz crystals, symmetrical in their outlines and in their relationship to the electrodes referred to a plane perpendicular to an electrical axis, are equivalent.

Accordingly, any piezoelectric oscillating quartz can always be uniquely determined by merely having its relationship to the r -face, r' -face and m -face clearly identified, and it does not matter whether the mother crystal is right-handed or left-handed.

LIX. *Notices respecting New Books.*

The Phase Rule and Its Applications. By ALEXANDER FINDLAY. Eighth edition. Revised with the assistance of A. N. CAMPBELL. [Pp. xv+327.] (London : Longmans, Green & Co., 1938. Price 12s. 6d.)

THE production of the eighth edition of this well-known book, thirty-four years after the first publication, is a splendid achievement, upon which Professor Findlay and his collaborator are to be heartily congratulated. Professor Findlay's concise style and his careful, deliberate treatment of each topic, which have materially contributed to the success of the book in the past, are still valuable features. Phases, components, degrees of freedom, and the phase rule expression are discussed in a short introductory chapter. Fourteen chapters are devoted to systems of one, two, and three components, whilst systems of four components are treated in a final chapter. A fair amount of modern work has been incorporated in the new edition, mainly in the form of additional references or short summaries, but the greater part of the material belongs, of necessity, to the classical period of physical chemistry. In certain cases, the modern aspect is perhaps not given sufficient prominence relative to the older work, as in the account of the influence of intensive drying on the inner equilibria of liquids. Typographical errors are extremely rare and the numerous diagrams are well drawn and clearly marked. The book is undoubtedly very good value at 12s. 6d., and should be included in every chemist's library.

R. S.

Modern Atomic Theory. By J. C. SPEAKMAN. [Pp. 208.] (London : Edward Arnold & Co., 1938. Price 6s.)

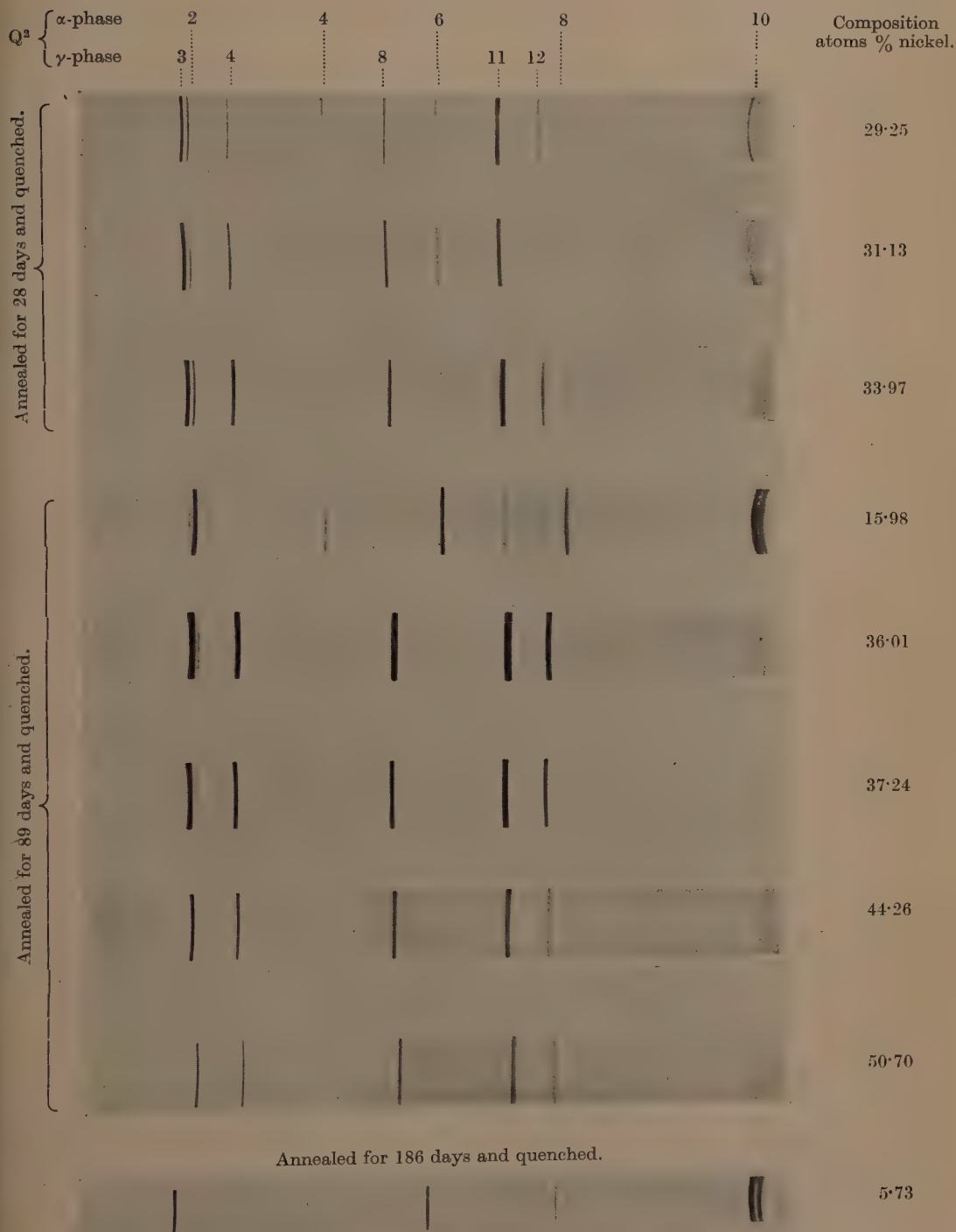
IN this survey of modern atomic theory the author has attempted to give, from an elementary standpoint, a logical presentation of the subject. "The human mind is so made," he states, "that it derives satisfaction from any hypothesis which can account for the observed variety [of properties of material things] on the basis of a few simple units of construction." Accordingly, after a preliminary chapter, the experimental evidence about the main "units of construction"—electrons, positrons, protons, alpha-particles, and neutrons—is described and discussed. An account follows of the particle properties of radiation and the wave properties of matter. On this basis are considered the extra-nuclear structure of the atom, mainly on Bohr theory lines, and then the composition and structure of the nucleus, there being good accounts of isotopy and of radioactivity.

Almost inevitably, theoretical conclusions must sometimes be stated without adequate indication of the evidence, particularly in connexion with those based on quantum mechanical considerations. In this respect, perhaps, the author has tried to include more than can be fitted into the ideal logical scheme. Usually, however, he has chosen for consideration those aspects of his subject which can, in fact, be satisfactorily treated in an essentially elementary and simple manner. The book is clearly and briskly written. It contains a number of excellent tables, and it is attractively produced. It may be recommended to students in universities (and possibly even in schools, as the author suggests) who require a short, well-proportioned and non-mathematical survey of the present outlook on atoms and their constituents.

E. C. S.

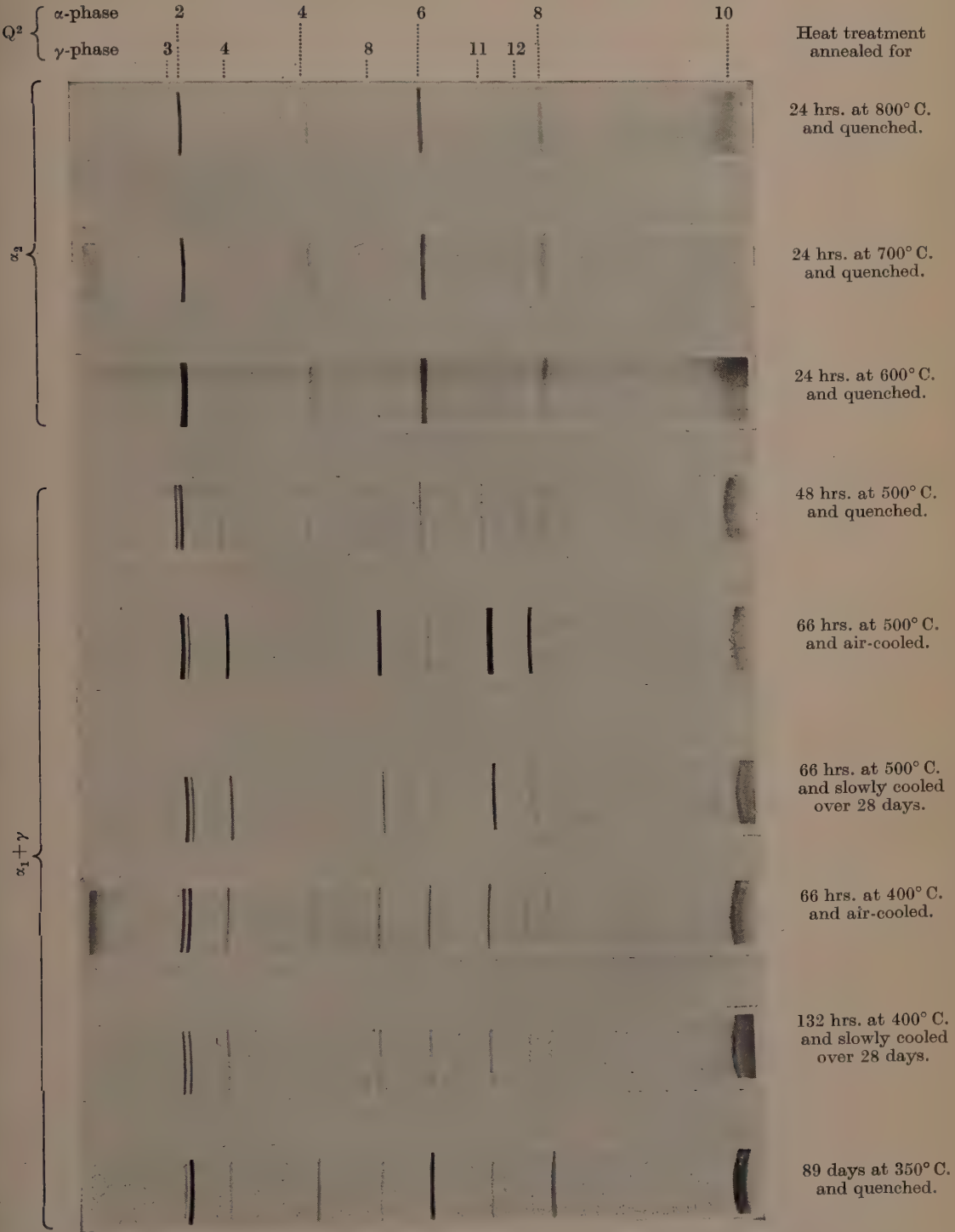
[The Editors do not hold themselves responsible for the views expressed by their correspondents.]

FIG. 3.



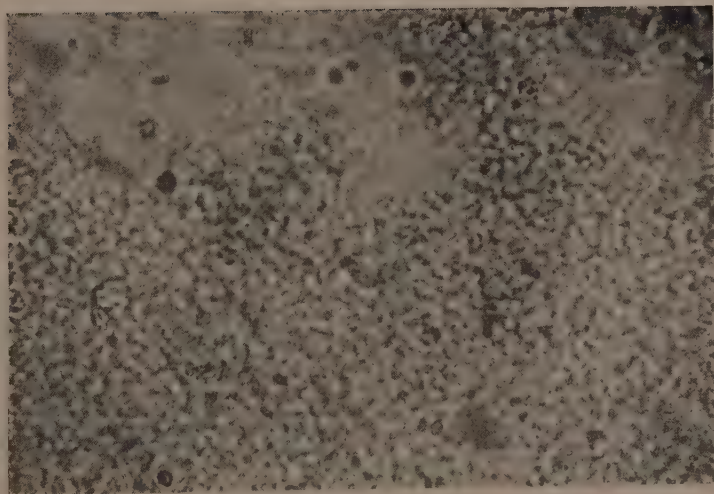
Iron-nickel alloys annealed at 350° C.

FIG. 4.



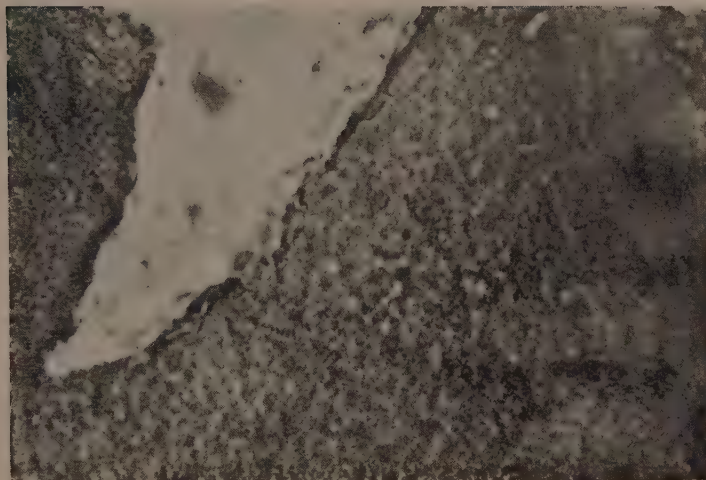
Spectra obtained with an iron-nickel alloy containing 16.0 atomic % nickel after different heat treatments.

FIG. 1.



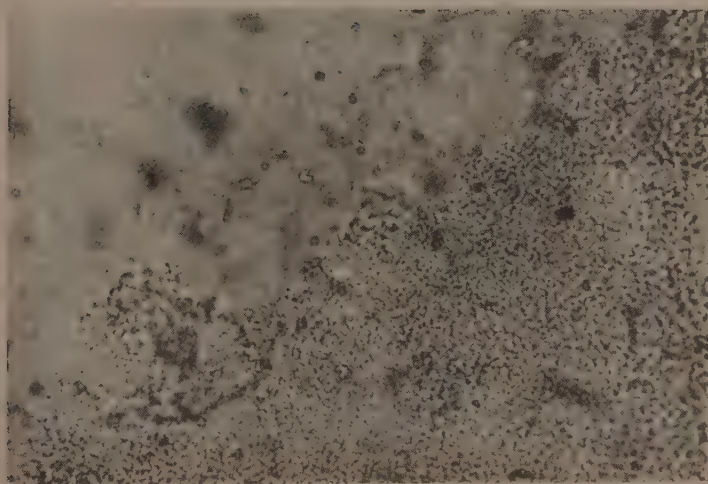
Blood-cells after pressure treatment at 5000 atmos. for $\frac{1}{2}$ hour. Magnification 1000 \times .

FIG. 2.



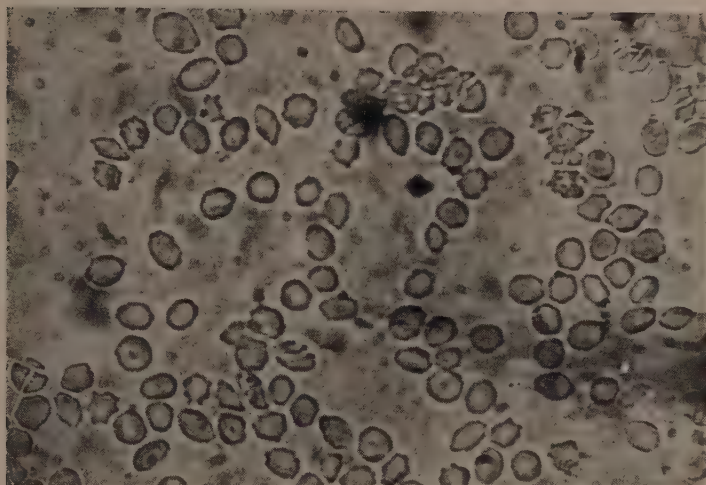
Aqueous solution of hemoglobin after pressure treatment at 5000 atmos. for 2 hours. Magnification 1000 \times .

FIG. 3.



Aqueous solution of hemoglobin after treatment with ethyl alcohol. Magnification 1000 \times .

FIG. 4.



Erythrocytes of normal bovine blood suspended in Hayem's solution. Magnification 1000 \times .

FIG. 5.



Erythrocytes disintegrated by pressure
treatment of 5000 atmos. for 1 hour.
Magnification $1000\times$.

LX. *On the General Validity of Nyquist's Theorem.*

By D. A. BELL, B.A., B.Sc.*

[Received February 27, 1939.]

Nyquist's theorem is thoroughly established already for simple ohmic resistors, and the recent work of F. C. Williams has dealt with complex circuits built up of a number of linear resistances and reactances. The present paper is concerned with the general basis of the theorem, and its application to circuit elements other than simple resistances and reactances.

NYQUIST'S theorem may be taken to be the statement that "Every electrical circuit element which can be represented as equivalent to a simple ohmic resistance is inherently the source of a random voltage, having components uniformly distributed over the observable frequency range, whose mean square magnitudes are linearly proportional to magnitude and temperature of resistance, with constant of proportionality such that the numerical relationship is

$$\bar{v}_{df}^2 = 4RkT \cdot df, \quad (1)$$

where \bar{v}_{df}^2 is mean square voltage in frequency band df , k is Boltzmann's constant, and T is temperature in degrees Kelvin or absolute."

There are several points to be considered in this apparently simple statement. First, it must be noted that we can discuss only a time-average mean-square value of voltage, and it represents a voltage not at a unique frequency, but summed over a narrow band of frequencies df . It is known, both empirically and theoretically, from the work of Rowland on shot noise †, that the average will, in fact, tend to a definite limiting value as the time increases; but the numerical length of time required for a given value of probable error does not appear to be known. If we show \bar{v}_{df}^2 explicitly as an integral over the frequency band df , equation (1) becomes

$$\bar{v}_{df}^2 = \int_f^{f+df} \bar{v}_f^2 \cdot df = 4RkT \cdot df. \quad (2)$$

The symbol \bar{v}^2 of course conceals a time integral, since the time average

* Communicated by Prof. E. V. Appleton, F.R.S.

† Bibliography (4).

must be found by a process of integration, so that with t for time equation (2) is in full

$$\bar{v}_{df}^2 = \frac{1}{dt} \int_f^{f+df} \int_t^{t+dt} v^2 \cdot df \cdot dt = 4RTk \cdot df, \quad \dots \quad (3)$$

or, in symmetrical form,

$$\int_f^{f+df} \int_t^{t+dt} v^2 \cdot df \cdot dt = 4RkT \cdot df \cdot dt. \quad \dots \quad (3a)$$

From this it appears that the reliability of a value of \bar{v}_{df}^2 so derived is proportional to the product $df \cdot dt$. This leads to the conclusion that in noise measurements the narrower the band width of the amplifier the longer must be the time-constant of the integrating device; in practice, of course, the time-constant is usually of very ample length for all practical requirements. This conclusion is an example of the general principle that in any statistical problem the relative fluctuations are inversely proportional to the number of constituents involved; if we reduce the number of frequency components involved, a given consistency of behaviour of \bar{v}^2 can only be retained by increasing the length of time over which the average is taken.

Having now decided exactly what is meant by \bar{v}_{df}^2 , and what are the inherent limitations to the accuracy of its measurement, it is convenient to divide the theorem into three parts:—

Part A.—“Equal resistances at equal temperatures have equal mean square fluctuation voltages, similarly distributed over all possible frequencies.”

Part B.—“The relation between magnitudes of resistance, temperature, frequency band, and mean square voltage is given by the equation $\bar{v}_{df}^2 = 4RkT \cdot df$.” We shall later add a new clause to Nyquist’s theorem, namely:—

Part C.—“In the general case of non-linear circuit elements whose current-voltage characteristics are of the type $I = aV^n$, the mean square fluctuation voltage is

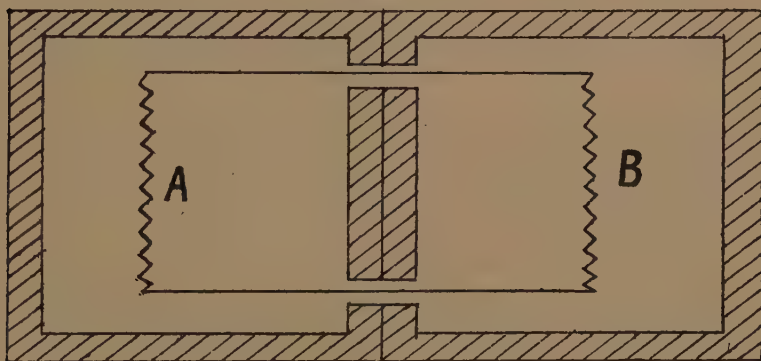
$$\bar{v}_{df}^2 = 4R_s kT \cdot df, \quad \dots \quad (1a)$$

where R_s is the differential resistance dV/dI of the non-linear circuit element at the operating point under consideration.”

Theorem A can be discussed formally and without considering the mechanism by which the noise originates, by a slight modification of Nyquist’s original treatment. It is necessary, however, to consider critically *every assumption* that is made in constructing the model to be used. Suppose we are given two resistances of equal value, but dissimilar construction, at the same temperature; these are to be electrically connected together to form a complete circuit, but isolated both from each

other and from all external surroundings for all forms of energy other than that conveyed by an electric current. Thus in fig. 1, the inner faces of the containing box must be assumed perfectly reflecting to all kinds of electromagnetic radiation (including heat, light, etc.), the interior would presumably be evacuated, and the resistances supported by threads of negligible conductivity for thermal and mechanical energy. Although perfection is not practically attainable, it is reasonable to assume that the enclosing walls could be made so good that the energy, of any type, passing through them would be negligible compared with the electrical energy transferred along the conductors joining the two resistances. The thermal conductivity of the electrical conductors joining the two resistances is a more serious difficulty. Change of cross-section alone will not help, for both thermal and electrical conductivity

Fig. 1.



are related to the geometry of the conductor in the same way. Neither does change of material help, if the choice be confined to metals: it is well known that different pure metals have similar values of the ratio of thermal to electrical conductivities, and alloys tend to have poor electrical conductivity. However, we may assume the use of a pair of ideal transformers, one at each resistance, giving very large step-up from resistance to line; the conductors may then be reduced in cross-section until the thermal conductivity is small, the electrical losses remaining small if the transformer ratio is high enough to give a very small current for a reasonable amount of power, and in the limiting case the thermal energy transferred can be made negligible compared with the electrical energy.

To the engineer this demand for "ideal" conditions which can never be attained in practice might appear almost sufficient weakness to invalidate the argument; but it is the method which has been employed in thermodynamics with complete success, and in reality it only implies

the assumption of uniformity of behaviour of the materials under investigation. This is a philosophical point which the author has not seen discussed. But for practical purposes it is an acceptable assumption that our resistance, for example, would inherently behave in the same way in a chamber giving perfect thermal insulation as in a chamber which is only very nearly perfect; and that the error due to imperfections is a continuous function of the magnitude of the imperfections, and can therefore be made as small as desired, though never completely eliminated, by taking sufficient precautions to reduce imperfections. Granted this assumption, we are justified in arguing from the ideal case.

Returning to fig. 1, we now assume that one of the resistances, A, is the source of a random e.m.f. of unknown mean square magnitude \bar{v}^2 . (This is the mean square value of the voltage summed over the whole frequency band from zero to infinity, not over a small range of frequencies only.) This assumption is experimentally justified provided that resistance A is of one of the known types, such as metal, carbon, or electrolytic resistance, which have been found to be the source of such an e.m.f.; but no conditions are attached to resistance B, save that it shall have the same value of resistance and temperature as A, and it may be of some new structure which is not experimentally known to be the source of a random e.m.f. The e.m.f. \bar{v}^2 will cause some current \bar{I}^2 , of unknown magnitude, to flow through the circuit, with liberation in each resistance of thermal energy $\bar{I}^2 R$. Now since both resistances are completely isolated from their surroundings, the energy required for the e.m.f. generated in A can only come from its store of thermal energy. (Apart from experimental and theoretical evidence of this, the only alternative is atomic disintegration, which could not continue indefinitely.) It follows, then, that thermal energy in A is constantly being converted into electrical energy, of which only half is restored to thermal energy in A, the other half reverting to thermal energy in B. Unless there is an exactly equal transfer of energy from B to A, the temperature of B will steadily rise and that of A will fall, which is contrary to the second law of thermodynamics and therefore regarded as impossible*. But under the conditions postulated for fig. 1, the only energy transfer is electrical, so that B must, in fact, be the source of an e.m.f. of equal mean square magnitude to that of A. Moreover, since it is possible to insert various types of frequency-selective circuits between A and B, and the insertion

* If in a self-acting system such a difference of temperature were continuously built up, energy could be obtained for nothing by connecting a heat engine between the two points at different temperature; for example, a steam engine could be used, with its boiler at the higher temperature and condenser at the lower.

of any such filter must not destroy the thermal equilibrium, it follows that the frequency-distribution of energy must be identical for the two resistances.

Part A of the theorem has thus been proved ; and its importance is that it enables the noise of *any* circuit element to be calculated, regardless of its internal structure, provided only that it has a determinable value of resistance and temperature. It should be noted that the proof is independent of any assumption as to the relation between mean square terminal p.d. of a resistance and the magnitude of external load shunted across the resistance, for a purely arbitrary value \bar{I}^2 was assumed for the current in the circuit. Neither is it really dependent upon the use of a particular resistance/power relationship, such as I^2R or V^2/R , and, consequently, it is not limited to linear resistances ; but the terminal p.d. of a non-linear resistance R' when connected to a linear resistance R must always be such as to dissipate in R an amount of energy equal to that which the terminal p.d. of R dissipates in R' . It does, however, require the assumption of independence of the two fluctuation e.m.f.'s, in support of which there are several arguments :—

(i.) There is overwhelming experimental evidence that the nature of the spontaneous e.m.f. generated in a resistance is "random," which by definition implies independence of external influence.

(ii.) The formulæ resulting from the hypothesis of independence have been verified. (See, for example, Bibliography (3).)

(iii.) Where information is available concerning the mechanism of the generation of the fluctuation noise, it is found to be of a type which is theoretically independent of external influence. (See, for example, the treatment of the metallic resistance in section 8 of Bibliography (1).)

So far, we have discussed only pure resistances, but the extension to linear complex networks has been thoroughly investigated both experimentally and theoretically by Dr. F. C. Williams (Bibliography (2)). The author considers that this can be extended to non-linear circuits after the manner explained above ; see discussion on Dr. Williams's paper, *Journ. I. E. E.* vol. lxxxiii. p. 432 (1938).

In a previous discussion on the treatment of thermionic valves as non-linear resistances, however, the author suggested the use of a correction factor depending upon the laws of curvature of the non-linear device. Dr. F. C. Williams has since pointed out that the derivation then used (Bibliography (1)) was incorrect, since by simple differentiation it employed the first powers of the small superimposed voltages ; for fluctuation voltages the first power terms must have mean value zero, only the squared terms being significant *. The treatment suggested as an alternative

* Dr. Williams's argument is published on p. 245 of Mr. E. B. Moullin's book 'Spontaneous Fluctuations of Voltage,' Oxford, 1938.

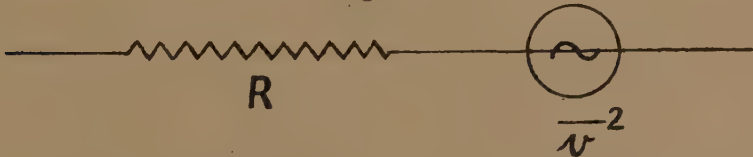
Part B of the theorem, being quantitative, is naturally dependent largely on direct experimental evidence. If one assumes that the equivalent circuit for a simple resistance is as shown in fig. 2, which is abundantly supported by experiment, then the proportionality between resistance and mean square voltage can be deduced by putting two resistances of unequal value, R_a and R_b , with mean square voltages \bar{v}_a^2 and \bar{v}_b^2 , in the apparatus of fig. 1. We then have two currents flowing, $\bar{i}_a^2 = \bar{v}_a^2 / (R_a + R_b)^2$, derived from the heat content of A, and the corresponding current $\bar{i}_b^2 = \bar{v}_b^2 / (R_a + R_b)^2$ derived from the heat content of B. Since each dissipates energy in the opposite resistance (as well as in its own resistance) at a rate $\bar{i}^2 R$, we have power transfer from A to B of

$$P_{ab} = \bar{i}_a^2 R_b = \bar{v}_a^2 R_b / (R_a + R_b)^2, \quad \dots \dots \dots (9)$$

and from B to A,

$$P_{ba} = \bar{i}_b^2 R_a = \bar{v}_b^2 R_a / (R_a + R_b)^2. \quad \dots \dots \dots (9a)$$

Fig. 2.



But for thermal equilibrium these two must be equal, which demands

$$\bar{v}_a^2 / R_a = \bar{v}_b^2 / R_b, \quad \dots \dots \dots (10)$$

i. e., \bar{v}^2 must be simply proportional to R .

It is possible that some thermal-electrical cycle could be devised to demonstrate that the mean square fluctuation voltage must be linearly proportional to the temperature, but the author has not been able to discover one. The linear variation with temperature, together with the constant of proportionality, can therefore only be demonstrated by the following methods :—

(1) Nyquist's original proof (Bibliography (5)) was similar in many respects to the proof of part A given above, but assumed a very long loss-free transmission line as the connecting link between the two resistances; the principle of equipartition of energy was then invoked by saying that the energy in the line must be such as to give the equipartition value $\frac{1}{2}kT$ per degree of freedom if the terminating resistances are removed from the line. It is necessary to adopt some additional hypothesis (e. g., quantum restrictions) to limit the energy attributed to very high frequencies, otherwise the total energy becomes infinite on taking the full frequency range up to infinity. This method has the advantage of being independent of any knowledge of the structure of the resistances. On the

other hand, it appears improbable that an entirely loss-free transmission line would contain spontaneous fluctuation voltages and currents; it appears that by virtue of the radiating properties of aërials, there is in practice a close parallel to the ideal loss-free transmission line, in the form of free space, which probably must be capable of being in equilibrium with large numbers of different circuits, at different temperatures, at the same time. A further difficulty is that if the hypothetical line is of *infinite* length there will be no reflexion to cause standing waves, and therefore no definable "degrees of freedom"; if, on the other hand, it is of finite length, there must be some lower limit of frequency below which the possible modes will not be sufficiently close together in frequency to cover all possible frequencies. This is probably not a sound objection, however, since it can always be specified that the length of the line is to be finite, but very great compared with any given wave-length which may be named as the lower limit of the frequency band under consideration.

(2) The fluctuation noise in a metallic resistance has been calculated from a theory of electronic conduction (see Bibliography (1)); by virtue of Part A of the theorem, this value may then be used for any conductor, metallic or otherwise. The chief weakness of this method is its use of a classical, instead of quantum, theory of electrical conduction*; but this defect may be mitigated to some extent by invoking the "correspondence principle," which states that the classical theory is always a limiting case to which the quantum theory approximates under suitable conditions. It has the advantage of showing directly the upper limit of frequency at which the theorem ceases to be applicable (namely, frequency such that the periodic time is comparable with the time of flight of the electrons within the conductor).

(3) There have been a number of experimental verifications of the results of the theorem, of which a recent example is given in (Bibliography (3)).

Thus we may summarize by saying that Part A can be put on a good philosophical basis, demanding only that there shall be continuous functions relating potential difference, current, and power; but Part B, on the other hand, rests largely on an empirical basis. This is not a serious limitation, since it is always possible to bring the problem of any unknown type of circuit element within the scope of Part A by supposing it to be balanced against an equal resistor of a type which has actually been experimentally examined. It has already been remarked that for this reason the author believes that Nyquist's theorem can be extended

* A better treatment has recently been given by Bakker and Heller; see Bibliography (9).

to cover non-linear circuit elements, using the expression given in equation (1 *a*).

This equation (1 *a*) points the way to an important unification of the whole field of fluctuation noise; for it will now be shown that with its aid one can suggest that there is in reality an intelligible relationship between the one extreme of the temperature-limited thermionic diode, the intermediate stage of space-charge-limited systems, and the other extreme of the simple metallic resistance. Simple shot noise theory shows that if the passage of the electrons conveying current between the terminals of a circuit element is independent both of the individual initial energies of the electrons and of the voltage applied across the terminals, as in the case of the temperature-limited diode or of the photo-cell, there is a mean-square equivalent fluctuation current of

$$\bar{I}_{sh}^2 = 2ie \cdot df, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

where i is the mean current flowing in the circuit. But if the circuit element has a finite resistance, though it need not be linear, we may employ the thermal theory to calculate a fluctuation current

$$\bar{I}_{th}^2 = (4kT/R_s)df \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

The resistance R_s is the effective value such that the energy dissipated in the circuit by mean square fluctuation voltage \bar{V}^2 is equal to V^2/R_s ; the mean square current corresponding to the mean square voltage $V^2 = 4R_s kT \cdot df$ has then been obtained by dividing the latter by R_s^2 , since R_s is the differential or slope resistance of the circuit element at the operating point in question. On dividing the value of mean square fluctuation current predicted by the thermal theory by the value which would be obtained from an equal mean current in conditions where the shot equation (11) is applicable, we obtain a factor which will be called the "mean square noise ratio" and denoted by F^2 ,

$$F^2 = \frac{\bar{I}_{th}^2}{\bar{I}_{sh}^2} = \frac{4kT}{2ieR_s} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

Now for a circuit element with law $i = aV^n$,

$$\frac{1}{R} = \frac{di}{dV} = naV^{n-1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

On substituting this value of R_s in equation (13),

$$F^2 = \frac{4kTn}{2eV} = n \cdot \frac{\frac{1}{2}kT}{\frac{1}{4}eV} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

The expression on the right-hand side of equation (15) consists of the

product of two factors whose qualitative significance has a very simple physical interpretation :—

(i.) The factor n is a function of the current-voltage law of the conductor, *i. e.*, of the ease with which the electron movements are influenced by the external voltage.

(ii.) The term $\frac{1}{2}kT/\frac{1}{4} eV$ is the ratio of the mean thermal kinetic energy of an electron (in equilibrium within the conductor at temperature T) measured in the direction of the external field, to the mean kinetic energy which the electron would have during an uninterrupted journey from terminal to terminal under a uniform acceleration due to the external potential difference. Note that the factor $\frac{1}{4}$ attached to eV arises because eV represents a *terminal* value of energy after a period of acceleration, which must be reduced to a mean value for the whole transit before it can be compared with the mean energy $\frac{1}{2}kT$.

At this point it may be as well to consider how the fluctuation current set up in a resistance can produce the fluctuation voltage which is observed in any measurement ; for according to the mechanism set out in Bibliography (1) the fluctuation *current* is the fundamental phenomenon. The explanation appears to be that any resistance must be shunted at least by its own self-capacitance, even though it may be so small as to have no perceptible effect on the impedance of the combination at the frequencies of measurement, however high in the radio-frequency scale. Every pulse of fluctuation current then charges this capacity, which subsequently discharges through the resistance itself and any other circuit elements which may be connected in parallel with it. It is for this reason that although the fluctuation *current* is the fundamental phenomenon, yet the sharing of energy between the resistance itself and any other circuit elements connected to it may be considered in terms of fluctuation voltage and factors of the type V^2/R .

We shall now examine two particular problems in which the nature of the “resistance” involved introduces sufficient difficulty for the question to be interesting and instructive.

First is the case of the thermionic valve, which will fall within the scope of Part C of the theorem, non-linear resistances, provided only the *temperature* of the valve resistance can be determined. Now the phenomenon of fluctuation noise in resistances evidently depends upon the fact that an electron, having both mass and charge, serves as a link between the thermal system, which might almost be described as a mechanical system, and the electrical system. The energy expended or absorbed by the electrons in reaction with an applied field therefore depends upon the vector products of the movements of the electrons and the field existing during their movements ; and in any conductor having definite

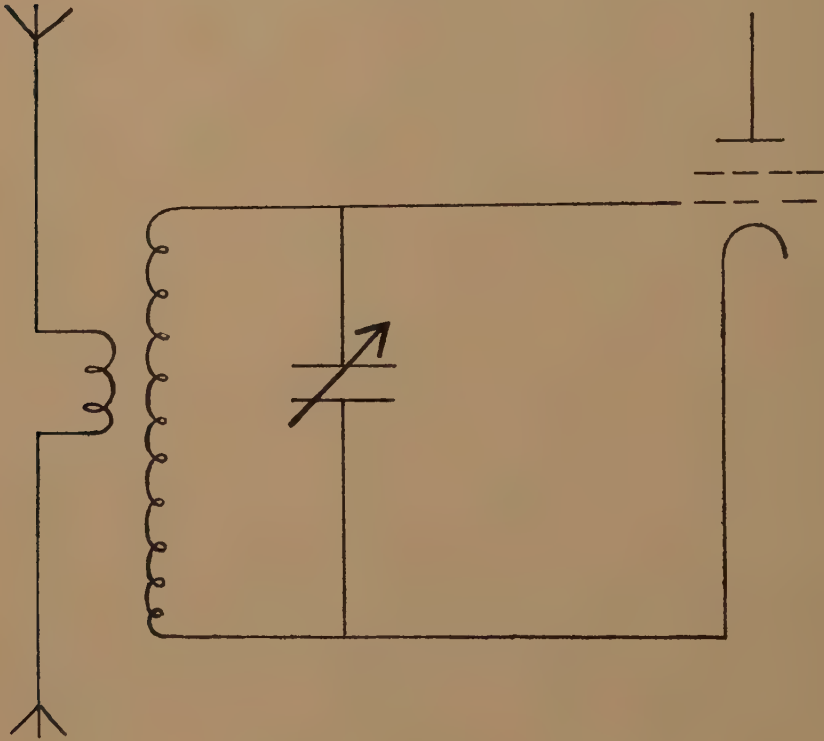
terminals to which external potentials are applied, including fluctuation potentials developed in other circuit elements, the field defines a fixed direction through every point of the conductor. The electron velocities may therefore be resolved into three Cartesian components, one along the direction of the field and the other two at right angles; the component along the direction of the field then enters into the energy exchanges, but the two components at right angles to the field give zero vector products with the field, and therefore have no effect on the energy exchanges. In a metallic conductor this consideration is of no importance, since the electrons change their directions of motion so frequently by collision with atoms that the average velocity will be the same along all directions; but in a thermionic valve this is not so, and it is necessary to work in terms of the components of velocity in the direction of the current flow only. It can then be found, as in Bibliography (1) and (6) for example, that the velocities of the electrons in the direction of the current flow are equivalent to a temperature of the order of half the cathode temperature. It appears to the author that to doubt the existence of thermal noise in the valve's internal resistance is then equivalent to doubting the universality of Part A of the theorem, which is otherwise unquestioned.

In the case of triodes and other amplifying valves, it should be remarked that the external circuit will feed energy into the valve *via* the anode, so that a treatment in terms of thermal noise as required by Nyquist's theorem must be based on the slope resistance of the path from anode to cathode within the valve. For this reason the author views with some suspicion all arguments which suggest that noise is "generated" in a localized part of the space-charge near the cathode, and amplified or smoothed, as the case may be, in some other part of the cathode-anode path. This difficulty applies to F.C. Williams's interpretation of Schottky's expression for the diode smoothing factor (quoted on pp. 223-4 of "Spontaneous Fluctuations of Voltage") to Schottky's treatment of the noise in triodes in terms of the resistance from cathode to grid plane (Bibliography (7)), and to the argument which led Percival and Horwood to direct their experimental investigations to triodes connected in circuit so as to be equivalent to diodes for alternating potentials, on the grounds that since the observed noise is greater than that predicted by a simple thermal theory, "... it would therefore appear that in some way an amplifying valve magnifies the effective temperature. What is required is clearly the effective temperature before this magnification takes place." (Bibliography (8).)

Another problem, which has been brought to the author's notice by Mr. O. L. Ratsey, is the effect of the radiation resistance of an aerial coupled to a receiving circuit. In a network containing a number of

resistive elements, the total noise is the resultant of a number of components depending upon the magnitudes and temperatures of the various resistances. Consider now fig. 3, which represents a half-wave dipole aerial coupled to a tuned circuit and valve. The resistive component of the impedance of the system, as measured at the points of connexion to the valve, must be partly due to aerial resistance, and with such an aerial it is easy to make the "radiation resistance" of the aerial much

Fig. 3.



greater than the copper and dielectric loss resistances ; what, then, is the temperature of this " radiation resistance " ?

It is easiest to consider the problem first in a very general way. The radiation resistance of an aerial can be calculated on the assumption that the aerial is situated in free space and is not influenced by any external conductors, dielectrics, etc., and it is not necessary for the purpose of calculation that it should have any internal loss ; but in that case its radiation is propagated through free space, and no mechanism can be found for the loss corresponding to the measured " radiation resistance."

In practice, of course, the radiation is gradually attenuated by losses in the earth, ionized regions of the atmosphere, receiving aerials, and, in fact, all matter in the radiation field; but none the less the "radiation resistance" is not seriously affected by these sources of dissipation, provided they are more than a wave-length or two distant, so that the practical aerial cannot behave in a manner very different from that of the ideal aerial in free space. It would then appear that we have a resistance which does not involve dissipation of electrical energy, and this is impossible to reconcile with the normal concept of resistance.

Now the usual method of calculating the "radiation resistance" of an aerial is to assume some *current* distribution which satisfies the conditions imposed by the distributed capacity of the aerial, and determine the e.m.f. which must accompany it, dividing the e.m.f. into in-phase and quadrature components; the ratio of in-phase e.m.f. to current is then called the "radiation resistance," while the ratio of quadrature component to current is called the reactance of the aerial. (See, for example, 2nd edition of 'Radio Frequency Measurements,' E. B. Moullin, London, 1931, p. 34; the ratio of in-phase components is here carefully described as "the apparent effective resistance, or 'radiation resistance.'") Since in the ideal case of an aerial in free space there is no dissipative agency to involve degradation of energy, but merely a reduction of energy density as the radiation spreads away from the aerial, it is difficult to see how there can be a true resistance; further, since only free space is involved, it is impossible to assign to it a temperature, for temperature is undoubtedly a property of matter*.

It is therefore suggested that, from the point of view of noise, an aerial with true resistance r and "radiation resistance" R' will behave in a similar manner to a resistance r in series with a reactance X numerically equal to R' (figs. 4 *a* and 4 *b*). It has already been shown that the "radiation resistance" is not a true dissipative resistance in the usual sense, and therefore presumably not a source of fluctuation noise; but if the equivalence suggested in fig. 4 *b* is to hold, it must also be shown that fluctuation current generated by another resistance which may be connected to the aerial, *e. g.*, a tuned circuit, will not dissipate energy in the "radiation resistance."

Although surprising at first sight, this second condition is quite possible. For "radiation resistance" depends upon a certain phase relationship between the electro-magnetic field and current, and is calculated for the case of a cyclic current of the frequency in question; but however much one may narrow the band-width df of a fluctuation current, it remains random, consisting of components derived from a number of random

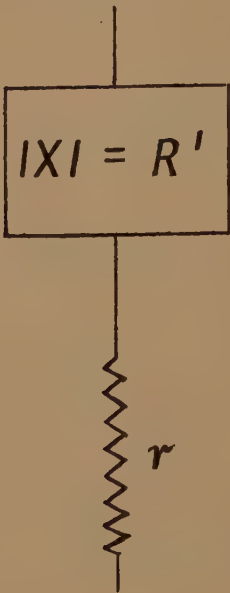
* "Temperature" may be defined as a function of the mean kinetic energy of random motion of any given collection of particles.

events, whose duration is extremely short compared with the periodic time of the current under consideration, and whose time-distribution is random. Consequently, the fluctuation current at any given frequency is subject to random phase modulation, and if the variations in phase occur at sufficiently close intervals (and from the random nature of the origin of the current it is probable that they do) the requisite special phase relationship between field and current is destroyed : the current at any instant is as likely to be in opposition or in quadrature as in like phase with the field established by the preceding current element, so that on

Fig. 4 a.



Fig. 4 b.



the whole the work done by the current on the field is zero and no energy is radiated.

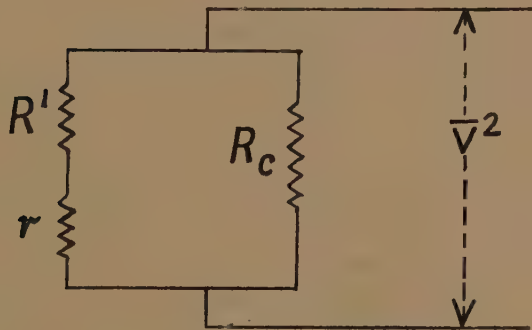
The equivalent circuit at resonance of the aerial of fig. 4 a coupled to a tuned circuit of dynamic resistance R_c is as shown in fig. 5, assuming that R' and r are now the effective values of the components of aerial resistance after transfer through the coupling system to the tuned circuit. The total mean square fluctuation voltage can then be found by applying Williams's theorem for a complex network (Bibliography (2)), which states that

$$\overline{v_{df}^2} = z_{AA}^2 \cdot 4k \, df \sum \frac{r_x T_x}{z_{Ax}^2}, \dots \dots \dots (16)$$

where \bar{v}_{df}^2 is the mean square fluctuation voltage in band df measured at terminals A, A'; $Z_{AA'}$ is modulus of network impedance between A and A'; Z_{Ax} is modulus of impedance from the hypothetical generator of fluctuation voltage in series with any given resistive element r_x and a short-circuit across A, A'; and r_x , T_x are the resistance and temperature of the given resistive element. Since we have suggested above that the "radiation resistance" does not behave as a resistance from the point of view of fluctuation currents, and that the e.m.f. across it is not in phase with the current, consistency demands that it should be treated as a reactance in evaluating both $Z_{AA'}^2$ and Z_{Ax}^2 . We then find for the squared impedance between A and A'

$$Z_{AA'}^2 = \frac{R^2(r^2 + R'^2)^2}{(r^2 + R'^2 + rR_c)^2 + R'^2 R_c^2},$$

Fig. 5.



and for the transfer impedances for aerial resistance r and circuit resistance R_c we find respectively

$$\begin{aligned} Z_{Ar}^2 &= r^2 + R'^2, \\ Z_{Ac}^2 &= R_c^2, \end{aligned}$$

so that equation (16) is evaluated as

$$\bar{v}_{df}^2 = 4k df \left\{ \frac{R_c^2(r^2 + R'^2)^2}{r(r^2 + R'^2 + rR_c)^2 + R'^2 R_c^2} \right\} \left\{ \frac{rT_r}{r^2 + R'^2} + \frac{T_c}{R_c} \right\}. \quad (17)$$

In the limiting case when r tends to zero compared with R' and R_c , equation (17) reduces to

$$\bar{v}_d^2 = 4k df \left\{ \frac{R_c^2}{1 + R_c^2/R'^2} \right\} \frac{T_c}{R_c}.$$

In the special case of $R_c = R'$, *i. e.*, correct impedance matching of aerial to circuit, this may be expressed as

$$\bar{v}_{df}^2 = 4\left(\frac{1}{2}R_c\right)kT_c df, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17a)$$

which is exactly the same as would be found if R' were a true resistance at the same temperature as R_c ; but if, as assumed above, R' is to be treated as a reactance, this correspondence will not hold for other ratios.

The above argument is probably open to criticism, but it is the only one of a number of possible hypotheses considered by the author which does not lead to the conclusion that a resistance connected to an aerial continually loses energy by electromagnetic radiation into space; if that occurred, the combination of aerial and resistance would presumably function as a self-acting refrigerator, which does not appear probable. An experimental test would be very difficult, owing to the stray signals, atmospherics, and other spasmodic noises likely to be picked up by an aerial at a strength comparable with the differences in fluctuation noise of the circuits predicted by various hypotheses.

The following are therefore the conclusions reached:—

(1) Nyquist's theorem can be applied to any network whose temperature and differential resistance are determinable. But a special case is the radiating aerial, whose "radiation resistance" is to be treated as equivalent to a reactance for the purpose of calculating fluctuation voltages.

(2) Equation (15) suggests a link between shot noise and thermal noise in terms of the respective mean energies of the electrons constituting the currents in the two cases.

Bibliography.

- (1) "A Theory of Fluctuation Noise," D. A. Bell, *Journ. I. E. E.* lxxxii. p. 522 (May 1938).
- (2) "Thermal Fluctuations in Complex Networks," F. C. Williams, *Journ. I. E. E.* lxxxi. p. 751 (Dec. 1937).
- (3) "Coexistent Thermal and Thermionic Fluctuations in Complex Networks," F. C. Williams, *Journ. I. E. E.* lxxxiii. p. 76 (July 1938).
- (4) "The Theory of the Mean Square Variations of a Function formed by adding Known Functions with Random Phases, and Applications to the Theories of the Shot Effect and of Light," E. N. Rowland, *Proc. Camb. Phil. Soc.* xxxii. p. 580 (1936).
- (5) "Thermal Agitation of Electric Charge in Conductors," H. Nyquist, 'Physical Review,' xxxii. p. 110 (1928).
- (6) "Effect of Space Charge and Transit Time on the Shot Noise in Diodes," A. J. Rack, *Bell Syst. Tech. Journ.* xvii. p. 592 (1938).
- (7) "Die Raumladungsschwächung des Schroteffektes," W. Schottky, *Wiss. Veröff. aus den Siemens-Werken*, xvi. part 2 (1937).
- (8) "Background Noise produced by Valves and Circuits," W. S. Percival and W. L. Horwood, 'Wireless Engineer,' xv. p. 128 (March 1938).
- (9) "On the Brownian Motion in Electric Resistances," C. J. Bakker and G. Heller, 'Physics,' vi. part 3, p. 262 (1939).

LXI. *The Dielectric Constant of Benzene.*

By WILFRED C. VAUGHAN, B.Sc., Ph.D.*

[Received February 14, 1939.]

THE determination of the dielectric constant of some standard liquid offers a very suitable means of judging the accuracy of any new method of comparing capacities. This investigation was carried out as a test of the author's original method described in a previous paper †. Benzene offers many unique advantages in this respect, and this liquid was chosen on account of its low dielectric constant and of the ease with which a satisfactorily pure specimen could be procured.

Many determinations of the dielectric constant of benzene have been carried out, although surprisingly few of these have been made at radio-frequencies. Examination of the available information reveals very good agreement with the value

$$\epsilon = 2.2825 \text{ at } 20^\circ \text{ C.}$$

for recent audio-frequency methods, but among results obtained at radio-frequencies the agreement is not so close. Among the latter the following results are of interest, each having been obtained in the region of 10^6 cycles/sec.

Worker.	$\epsilon_{25^\circ \text{ C.}}$	Reference.
Graffunder	2.268	<i>Ann. der Physik</i> , lxx. p. 225 (1923).
Isnardi	2.268	<i>Z. für Physik</i> , ix. p. 324 (1922).
Grutzmacher.....	2.277	<i>Z. für Physik</i> , xxviii. p. 343 (1924).
Sugden.....	2.271	<i>J. Chem. Soc.</i> (July 1933).

Resonance methods were employed in each of the above determinations, but unfortunately only in the last-mentioned reference is sufficient information given to enable the degree of accuracy to be judged. In this case a substitution method was used, the standard variable condenser having a maximum value of $1200 \mu\mu\text{F.}$ Its scale could be read to 0.02° by means of a magnifier, so that variations as small as about $0.1 \mu\mu\text{F.}$ could be observed. The test condenser had an air capacity of about $190 \mu\mu\text{F.}$, and since its capacity was determined as the difference of two scale readings on the standard variable, the possibility of error

* Communicated by Dr. S. Marsh.

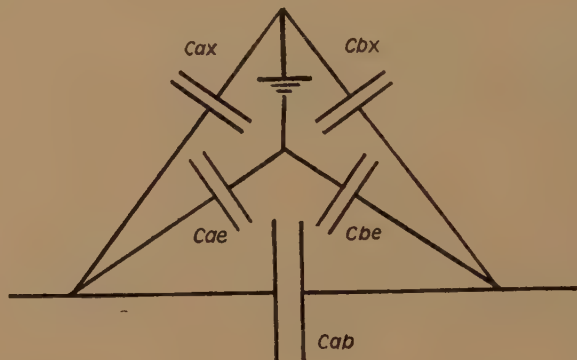
† Vaughan, *Phil. Mag.* xxvi. p. 521 (Oct. 1938).

was about $0.2\mu\mu\text{F.}$ or 0.16 per cent. When filled with benzene the capacity of the test condenser was approximately doubled, so that a further error of 0.08 per cent. might result. On the assumption that the accuracy of adjustment was considerably superior to that of reading the standard variable—and this appeared to be the case from the information given—the total error was probably in the region of 0.24 per cent. This might have been increased slightly by end effects and lead capacity and inductance, although it was pointed out that special consideration had been given to the reduction of errors of this nature.

Fig. 1.



Fig. 2.

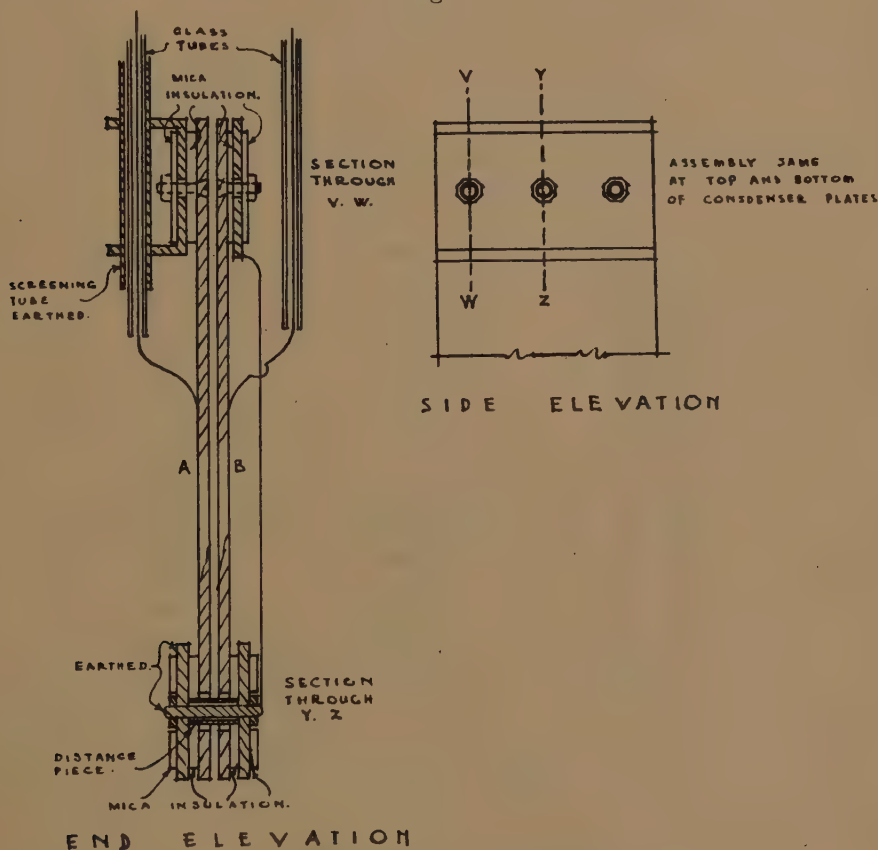


The present investigation involves the use of a test condenser which is first of all compared with a fixed standard. It is then filled with benzene, and the comparison repeated. In the previous paper it was shown unnecessary to take any account of stray capacities from the high potential side to earth, unless these were excessive. This allows the use of a foreign dielectric between the plates, providing that suitable earthed screens prevent lines of force from one plate reaching the other through the foreign dielectric. The design of a suitable test condenser is therefore greatly simplified.

Suppose A and B (fig. 1) are two such condenser plates separated by the insulating material X and an earthed screen E. The capacity to be measured is C_{ab} (fig. 2), but stray capacities C_{ae} and C_{be} (through air)

and C_{ax} and C_{bx} (through X) are also present. All these stray capacities are to earth and, as has already been pointed out, will have no effect on the determination of C_{ab} . On filling the condenser with a liquid of dielectric constant ϵ , the magnitude of the stray capacity will alter on account of that part of the field which reaches E through the liquid. This, however, will have no influence on the determination of C_{ab} , since it remains a capacity of earth.

Fig. 3.

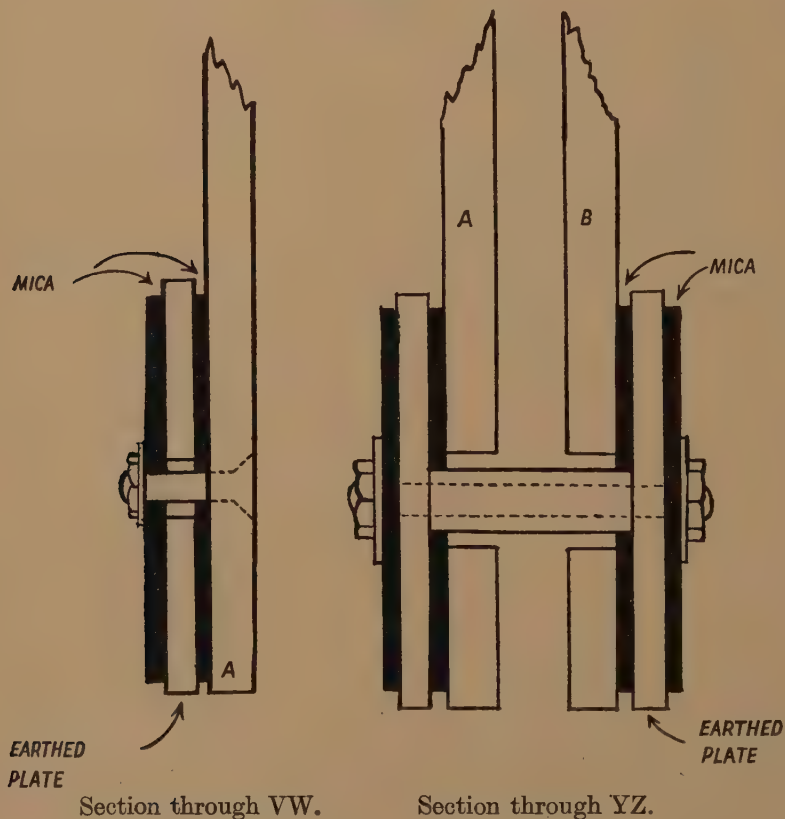


Construction of test condenser.

The test condenser was of very compact form and consisted of a pair of parallel brass plates supported on mica strips which were carried by earthed strips of brass. The method of assembly is shown in figs. 3 and 4. The separation of the plates A and B is maintained by two short lengths of brass tube which serve as distance pieces, one at the top and the other at the bottom of the assembly. It will be noticed that these butt

against the outer earthed brass strips, but do not touch either A or B. Connexion with the high potential plate was made by means of a lead which passed down the interior of an earthed copper tube from which it was insulated by a length of glass tubing. This copper tube formed a support for the condenser. The lead to the low potential plate was unscreened. Connexion between the oscillator and the high potential plate was made

Fig. 4.



by means of a length of screened cable. Using this arrangement all stray capacities were to earth, so that the method yielded the capacity between the parallel plates directly, without the necessity for troublesome corrections.

The condenser in its final form had a capacity of about $50 \mu\mu\text{F.}$ in air, and it was compared with a standard of about 1.3 times this capacity.

The test condenser was placed inside a large boiling-tube, the cork of which carried the whole assembly as well as a thermometer and two

glass tubes. One of the latter extended to the bottom of the tube. In order that the drying processes should be entirely satisfactory the tube and its contents were made completely air-tight by means of sealing-wax.

Preparation of the Benzene.

It has been pointed out by previous workers that considerable precautions are necessary to free the benzene from all traces of water. Preliminary experiments with the apparatus described, using benzene in the undried condition in which it was received, yielded results some 1 to 1.5 per cent. higher than those subsequently obtained after careful drying. In these investigations a reduction in the dielectric constant was obtained by dropping into the benzene some pieces of freshly cut sodium, and immediately sealing the tube. A slight action was observed, presumably due to the presence of water. This ceased after some hours, at the end of which time a result was observed only slightly in excess of that obtained after the more elaborate method of drying.

The sample of benzene used was supplied by the firm of May and Baker, Battersea, as pure recrystallizable benzene free from thiophene. This was allowed to stand over sodium wire for several weeks before use and was then distilled directly into the condenser vessel. The latter had previously been carefully dried both by exhausting and heating, and then by aspirating dry air through it for many hours prior to filling. In the filling process the escaping air passed through fused calcium chloride and strong sulphuric acid in order to prevent the reabsorption of moisture from the atmosphere.

Experimental Procedure.

The condenser vessel was placed in a vacuum flask containing paraffin which could be heated to any desired temperature. On account of the upper exposed metallic part of the condenser assembly it was difficult to maintain reasonably constant temperatures much above that of the room. A considerable improvement was made by covering the exposed parts with cotton wool.

All comparisons were made with a small semi-variable condenser which was originally set to about 1.3 times the capacity of the test condenser in air. This was enclosed in a copper screening box where it was left untouched until the completion of the investigation. The test and standard condensers being connected, the frequency of the oscillator was determined by means of the wave-meter. The comparison of the two condensers was then carried out exactly as described in the earlier paper. After this had been completed the test condenser was disconnected, carefully dried, and filled with benzene. It was then reconnected, and

the oscillatory circuit retuned to the original frequency by reducing the capacity of the tuning condenser after slightly reducing the capacity

TABLE I.

Comparison of Test Condenser in Air with Standard.

Frequency : 8.652×10^5 cycles per sec.

Temperature : 15°C .

Resistance in Series with Test Condenser C_t .

R, ohms.	log R.	P_t volts $\times 10^{-1}$.	log P_t .	log R_t/P_t .	log R_s/P_s .	log R_t/R_s .	log C_s/C_t .	C_s/C_t .
67.85	1.8315	0.7345	$\bar{2}.8660$	2.9655	2.7008	0.2647	0.1323	1.356
90.40	1.9562	0.9736	$\bar{2}.9884$	2.9678	2.7024	0.2654	0.1327	1.357
113.7	2.0557	1.214	$\bar{1}.0845$	2.9702	2.7061	0.2641	0.1320	1.355
135.1	2.1326	1.444	$\bar{1}.1596$	2.9730	2.7079	0.2651	0.1325	1.357
158.5	2.2001	1.678	$\bar{1}.2248$	2.9753	2.7098	0.2655	0.1327	1.357
181.4	2.2586	1.912	$\bar{1}.2814$	2.9772	2.7118	0.2654	0.1327	1.357
204.5	2.3107	2.145	$\bar{1}.3314$	2.9793	2.7138	0.2655	0.1327	1.357
227.8	2.3575	2.380	$\bar{1}.3766$	2.9809	2.7159	0.2650	0.1325	1.357

Resistance in Series with Standard Condenser C_s .

R, ohms.	log R.	P_s volts $\times 10^{-1}$.	log P_s .	log R_s/P_s .	log R_t/P_t .	log R_t/R_s .	log C_s/C_t .	C_s/C_t .
38.04	1.5803	0.7567	2.8789	2.7014	2.9657	0.2643	0.1321	1.355
50.92	1.7069	1.0062	$\bar{1}.0026$	2.7041	2.9682	0.2641	0.1320	1.355
63.78	1.8047	1.254	$\bar{1}.0983$	2.7064	2.9708	0.2644	0.1322	1.356
76.92	1.8860	1.504	$\bar{1}.1772$	2.7088	2.9734	0.2646	0.1323	1.356
89.90	1.9538	1.752	$\bar{1}.2435$	2.7103	2.9762	0.2659	0.1329	1.358
103.1	2.0133	1.998	$\bar{1}.3007$	2.7126	2.9780	0.2654	0.1327	1.357
116.3	2.0655	2.236	$\bar{1}.3495$	2.7160	2.9800	0.2640	0.1320	1.355
129.4	2.1119	2.477	$\bar{1}.3941$	2.7178	2.9818	0.2640	0.1320	1.355
Mean value of C_s/C_t (16 results)								1.3562

of the coupling condenser. After recomparing the capacities of test and standard condensers at three different temperatures, the vessel was emptied, thoroughly washed out with acetone, and carefully dried. This could easily be carried out without disturbing the assembly, and subsequent comparison with the standard gave ample proof of the rigidity of the test condenser.

Results.

The design of the condenser did not permit determinations at temperatures differing greatly from that of the room. It was possible, however,

TABLE II.

Comparison of Test Condenser filled with Benzene with Standard.

Frequency : 8.652×10^5 cycles per sec.

Temperature : 15.3°C .

Resistance in Series with Test Condenser ϵC_t .

R, ohms.	log R.	P _t volts.	log P _t .	log R _t /P _t .	log R _s /P _s .	log R _s /R _t .	log $\epsilon C_t/C_s$.	$\epsilon C_t/C_s$.
38.04	1.5804	0.2255	1.4074	2.1729	2.6307	0.4578	0.2289	1.693
50.92	1.7069	0.3433	1.5344	2.1723	2.6297	0.4574	0.2287	1.693
63.78	1.8047	0.4300	1.6335	2.1713	2.6286	0.4573	0.2286	1.692
76.92	1.8860	0.5202	1.7162	2.1698	2.6271	0.4573	0.2286	1.692

Resistance in Series with Standard Condenser C_t .

R, ohms.	log R.	P _t volts.	log P _s .	log R _s /P _s .	log R _t /P _t .	log R _s /R _t .	log $\epsilon C_t/C_s$.	$\epsilon C_t/C_s$.
113.7	2.0557	0.2661	1.4251	2.6305	2.1729	0.4577	0.2288	1.693
135.1	2.1326	0.3182	1.5027	2.6299	2.1725	0.4574	0.2287	1.693
158.5	2.2001	0.3721	1.5706	2.6295	2.1720	0.4575	0.2287	1.693
181.4	2.2586	0.4265	1.6299	2.6287	2.1713	0.4574	0.2287	1.693
204.5	2.3107	0.4820	1.6830	2.6277	2.1705	0.4572	0.2286	1.692
227.8	2.3575	0.5382	1.7310	2.6265	2.1695	0.4572	0.2286	1.692
Mean value of C_t/C_s (10 results).....								1.6926

$$\begin{aligned}\text{Dielectric constant of benzene} &= 1.6926 \times 1.3562 \\ &= 2.2951 \text{ at } 15.3^\circ \text{C}.\end{aligned}$$

to obtain reliable results at 15.3°C ., 19°C ., and 21°C ., without difficulty. The variation in the air capacity of the test condenser was found to be inappreciable over this range of temperature.

Tables I. and II. show a complete set of results and calculations relating to the observations at one of the above temperatures. A summary of all results is given in Table III.

TABLE III.

Dielectric Constant of Benzene.

Frequency : 8.652×10^5 cycles/sec.

Temperature.	C_s/C_t .	$\epsilon C_s/C_s$.	ϵ .
15.3° C.	1.3562	1.6926	2.2951
19.0° C.	1.3562	1.6875	2.2884
21.0° C.	1.3562	1.6804	2.2790
$\frac{\partial \epsilon}{\partial t}$ between 15.3° C. and 21.0° C. = -0.0028			

Assuming linearity, the value of $\frac{\partial \epsilon}{\partial t}$ between 15.3° C. and 21.0° C. is

$$\frac{\partial \epsilon}{\partial t} = -0.0028 \text{ per } ^\circ\text{C.}$$

This value yields the following interpolated and extrapolated values of the dielectric constant :—

Temperature.	ϵ .
20° C.	2.2818
25° C.	2.2678

It will be observed that these results are in satisfactory agreement with those quoted earlier in the paper.

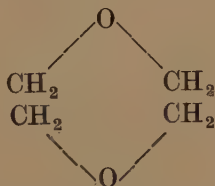
LXII. *The Dielectric Constant, Dipole Moment, and Molecular Polarization of 1.4 Dioxane (C₄H₈O₂).*

By WILFRED C. VAUGHAN, B.Sc., Ph.D.*

[Received February 14, 1939.]

THE use of dioxane as a solvent in connexion with the determination of dipole moments has engaged considerable attention, and has formed the subject of several recent papers. J. Warren Williams⁽¹⁾, Anschütz and Broeker⁽²⁾, and Reid and Hofmann⁽³⁾ have shown that this substance forms an excellent solvent for a considerable number of substances, inorganic as well as organic. Its low dielectric constant and almost negligible dipole moment make it a serious rival to benzene, which has hitherto been regarded as one of the best solvents for use in studies of electric moments.

The symmetrical form of the structural formula for 1.4 dioxane (diethylenedioxiide),



seems to indicate a zero value for the dipole moment, and J. Warren Williams⁽¹⁾ has shown that the value of μ cannot be greater than 0.4×10^{-18} E.S.U. This has been confirmed in the case of its vapour between temperatures of 64° C. and 112° C. by C. H. Schwingel and E. W. Greene⁽⁴⁾ from determinations of the dielectric constant.

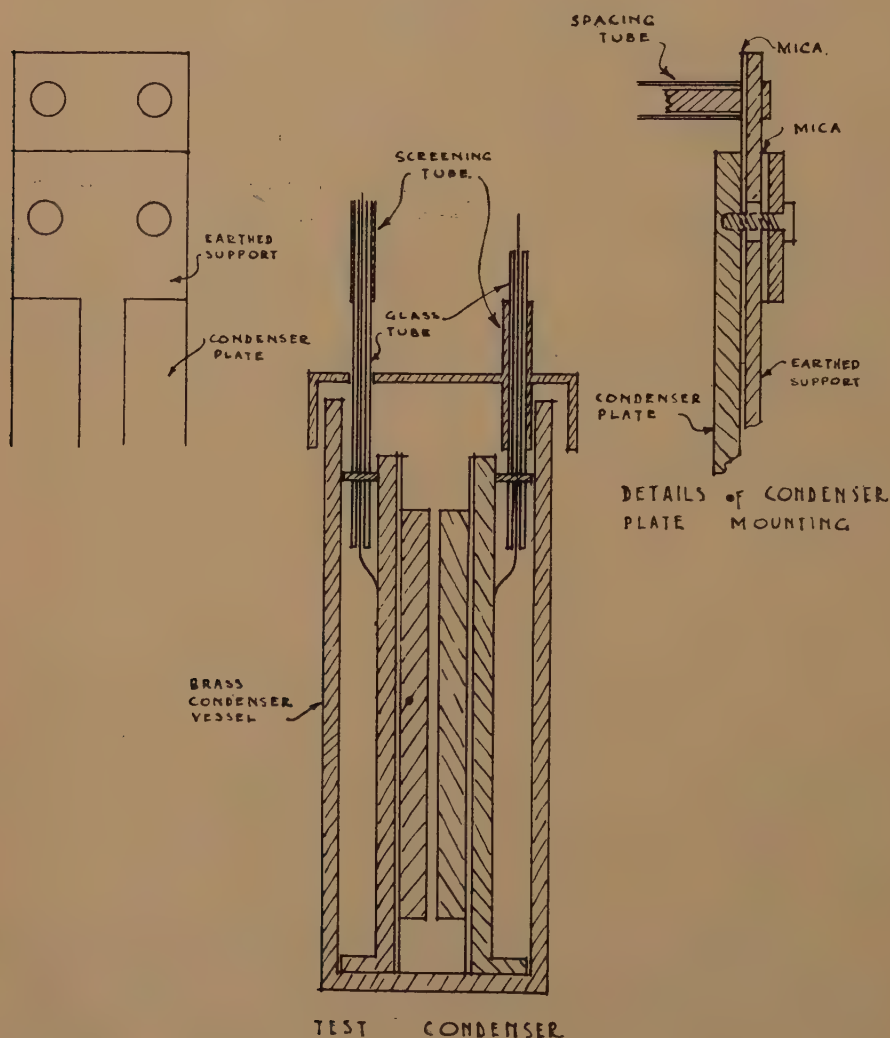
It was surprising to find from an examination of the available information that few accurate determinations of the dielectric constant of the pure solvent in the liquid form had been made, and, further, that there were apparently no data covering a range of temperatures. Such information would manifestly be of considerable importance, and it was therefore decided to carry out the following series of investigations using the method of measuring capacities which forms the subject of an earlier paper by the author*.

* Communicated by Dr. S. Marsh.

† Vaughan, Phil. Mag. xxvi. p. 521 (Oct. 1938).

The design and details of the test condenser used for this purpose are shown in fig. 1. The actual condenser plates were of brass $3/16''$ in thickness, and these were supported on mica bolted to earthed brass plates

Fig. 1.



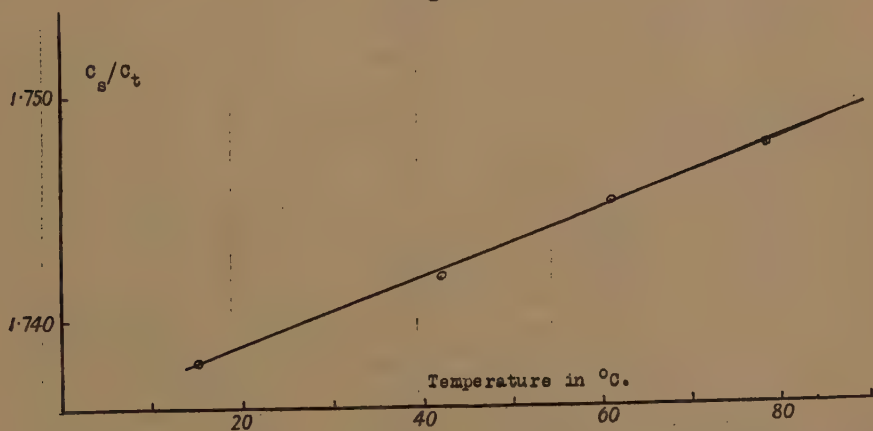
TEST CONDENSER

of the same thickness. The spacing was maintained by brass distance pieces between the earthed plates. This assembly fitted inside a brass tube on which a heating coil was wound. The condenser vessel was well lagged with asbestos, and was mounted on a flat steel plate from which it was thermally insulated by an asbestos mat. All leads passed through

small ebonite bushes set in the base plate. A large bell jar placed over the plate allowed the assembly to be exhausted after all joints had been sealed off with sealing wax. A pair of glass tubes and a thermometer were carried by a rubber stopper at the top of the bell jar. The tubes enabled the condenser to be filled and emptied without disturbing the assembly.

After assembling, the first process was the removal of all trace of moisture from the condenser and its associated mountings. The temperature was raised to about 200°C . and the pressure inside the bell jar maintained at about 1 mm. of mercury by means of a Hyvac pump. Following this, dry air was aspirated through the apparatus for many hours. Prolonged drying was necessary on account of the moisture absorbed by the asbestos. When this process had been completed, the capacity of the test condenser

Fig. 2.



was compared with that of a small standard which had been adjusted to about 1.7 of the capacity of the former. The method of comparison was identical with that described earlier in connexion with the dielectric constant of benzene*. Comparisons were made at 15°C ., 42°C ., 61°C ., 78°C ., and 91°C . As a precaution, the temperature was raised to about 180°C ., and then allowed to fall. Redeterminations at lower temperatures indicated that strains had not been set up by the process of heating and that results were repeatable. In all measurements the thermometer was situated inside the brass condenser vessel, and its reading was allowed to remain steady for about 5 minutes before making a determination. The results of these experiments, which are illustrated in fig. 2, showed practically a linear fall in capacity of the test condenser with rise in temperature.

* *Phil. Mag.* xxvii. p. 661 (1939).

The sample of 1·4 dioxane which formed the subject of this investigation was supplied by L. Light and Co. It was carefully purified both by distillation and by freezing, and determination of its melting- and freezing-points, as well as its refractive index before and after use, showed that it was extremely pure and had not been contaminated by the condenser vessel. The author is greatly indebted to Dr. J. Kenyon of the Chemistry Department, Battersea Polytechnic, who kindly undertook this part of the work.

After purification the dioxane was distilled directly into the condenser vessel which, since the earlier part of the experiment, had been sealed off to exclude all moisture. The substance under test was therefore in a very dry condition. After filling, comparisons were again made

TABLE I.

Test condenser in air.		Test condenser in dioxane.	
Temperature, °C.	$C_s/C.$	Temperature, °C.	$\epsilon C_t/C_s.$
15·0	1·7381	14·2	1·3431
42·0	1·7418	20·5	1·3164
61·2	1·7451	38·5	1·2930
78·8	1·7475	60·0	1·2809
91·0	1·7485	71·3	1·2740
		80·1	1·2691
		87·1	1·2661

between the test and standard condensers at various temperatures between 14° C. and 90° C. The results of all comparisons are shown in Table I. In conclusion, the test condenser was emptied, carefully washed out with acetone, dried, and again compared with the standard. This gave satisfactory evidence that the latter had undergone no change during the period covered by the experiments.

Results

From values interpolated from fig. 2. the dielectric constant was calculated at seven temperatures. The results are shown in Table II., whilst fig. 3 shows the variation of dielectric constant with temperature. From the latter the following results are interpolated :—

$$\begin{aligned}\epsilon \text{ at } 25^\circ \text{ C.} &= 2\cdot280, \\ \epsilon \text{ at } 50^\circ \text{ C.} &= 2\cdot242.\end{aligned}$$

Fig. 3.

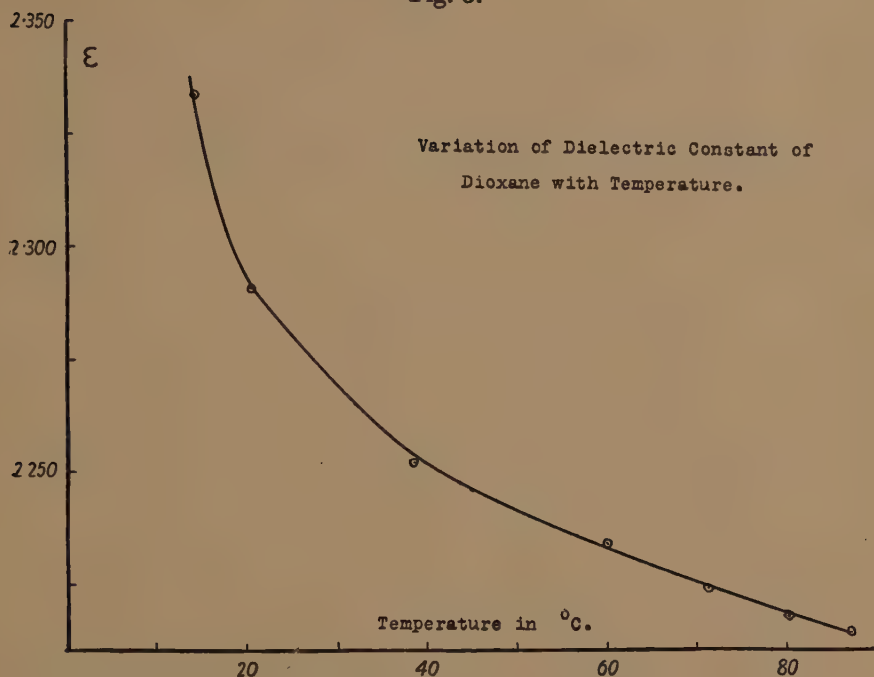


TABLE II.

Calculation of ϵ . Values of C_s/C_t interpolated from fig. 2.

Temperature.	$\epsilon C_t/C_s$.	C_s/C_t .	ϵ .
14.2	1.3431	1.738	2.334
20.5	1.3164	1.740	2.291
38.5	1.2930	1.741	2.252
60.0	1.2809	1.744	2.234
71.3	1.2740	1.746	2.224
80.1	1.2691	1.748	2.218
87.1	1.2661	1.749	2.214

The determination was made at a frequency of approximately 10^6 cycles/sec.

For purposes of comparison the results of other workers are given below :

Worker.	$\epsilon_{25^\circ \text{C.}}$	$\epsilon_{50^\circ \text{C.}}$
Smyth and Walls ⁽⁵⁾	2.306	2.251
Warren Williams ⁽¹⁾	2.2	
Kraus and Fuoss ⁽⁶⁾	2.2	

Discussion of Results.

The dielectric constant of dioxane is found to fall with increasing temperature, the curve of its variation being shown in fig. 3. From the point of view of its use in experiments on dipole moments it is interesting to note that the value of ϵ falls by only about 5 per cent. in a temperature rise of over 70° C. Unfortunately, the rate of decrease is greatest in the lower part of the temperature range.

TABLE III.

Molecular Polarization of 1.4 Dioxane.

Molecular Weight M of Dioxane ($C_4H_8O_2$) = 88.Clausius-Mossotti Factor $F = \frac{\epsilon - 1}{\epsilon + 2} \cdot \frac{1}{\rho}$.Molecular Polarization $P = M \times F$.Value of ϵ interpolated from fig. 3.

Temperature, °C.	ϵ .	ρ .	$\frac{\epsilon - 1}{\epsilon + 2}$	F.	P.
16.0	2.311	1.0377	0.3041	0.2930	25.78
20.0	2.294	1.0329	0.3013	0.2927	25.76
25.0	2.280	1.0311	0.2991	0.2908	25.59
50.0	2.242	1.0027	0.2928	0.2920	25.70
Mean value of P					25.71

$$\begin{aligned} \text{Polarization of molecule } \gamma &= \frac{3}{4\pi N} \times P. \\ &= 1.013 \times 10^{-23}. \end{aligned}$$

The Clausius-Mossotti factor,

$$\frac{(\epsilon - 1)}{(\epsilon + 2)} \cdot \frac{1}{\rho},$$

may now be calculated for the different temperatures covered by this investigation. As no means were available for the measurement of the density ρ , the values of previous determinations were taken. Results relating to temperatures of 25° C. and 50° C. are given in the paper of Smyth and Walls mentioned above, and further values for 16° C. and 20° C. were obtained from Beilstein's 'Handbuch der Organischen Chemie.' The value of the Clausius-Mosotti factor is shown in Table III. to be

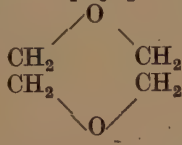
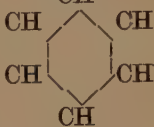
practically independent of temperature. Thus the term β in Debye's equation

$$\frac{(\epsilon-1)}{(\epsilon+2)} \cdot \frac{M}{\rho} = \alpha + \frac{\beta}{T}$$

is seen to be zero. If the above equation can be regarded as strictly applicable to a liquid then the dipole moment of 1·4 dioxane must be zero.

TABLE IV.

Comparison of 1·4 Dioxane and Benzene.

	1·4 Dioxane, $C_4H_8O_2$ 	Benzene, C_6H_6 
Molecular weight M	88	78
Density $\rho_{20^\circ C.}$ (water at $4^\circ C.$)	1·0329	0·8791
Melting-point	$9\cdot5^\circ C.$	$5\cdot5^\circ C.$
Boiling-point	$101\cdot0^\circ C.$	$79\cdot6^\circ C.$
Refractive index $n_D 20^\circ C.$	1·4165	1·5014
Dielectric constant $\epsilon_{20^\circ C.}$	2·294	2·282
Polarization P	25·71	26·69
Polarization of molecule γ	$1\cdot013 \times 10^{-23}$	$1\cdot052 \times 10^{-23}$

The molecular weight M of dioxane is 88, so that the molecular polarization P calculated from

$$P = \frac{(\epsilon-1)}{(\epsilon+2)} \cdot \frac{M}{\rho}$$

is 25·71. The actual significance of P is revealed in Debye's general equation

$$P = \frac{(\epsilon-1)}{(\epsilon+2)} \cdot \frac{M}{\rho} = \frac{4\pi N}{3} \cdot \gamma + \frac{4\pi N}{3} \cdot \frac{\mu^2}{3kT}$$

For non-polar liquids the final term, which is a function of temperature, will be zero. N is the number of molecules per gram-molecule ($6\cdot061 \times 10^{23}$) and γ is the molecular polarizability, being defined as the moment of the

doublet induced in a molecule by a field of unit strength. The value of γ has been calculated from

$$\gamma = \frac{3}{4\pi N} \frac{(\epsilon - 1)}{(\epsilon + 2)} \cdot \frac{M}{\rho},$$

using the results already obtained. Thus, for 1·4 dioxane γ is 1.013×10^{-23} .

In many respects 1·4 dioxane shows a marked resemblance to benzene as shown in Table IV. The similarity of the electrical properties ϵ , P , and γ for these two substances is of considerable interest.

References.

- (1) J. Warren Williams, "Use of Dioxane as a Solvent for Electric Moment Studies," J. Amer. Chem. Soc. lii. p. 1838 (1930).
- (2) Anschutz and Broeker, *Berichte*, lix. B, p. 1838 (1930).
- (3) Reid and Hofman, Ind. Chem. Eng. xxi. p. 695 (1929).
- (4) C. H. Schwingel and E. W. Greene, J. Amer. Chem. Soc. lvi. p. 683 (1934).
- (5) C. P. Smyth and W. S. Walls, J. Amer. Chem. Soc. liii. p. 2115 (1931).
- (6) Krauss and Fuoss, J. Amer. Chem. Soc. lv. p. 21 (1933).

LXIII. *The Thermal and Electrical Conductivities of some Magnesium Alloys.*

By R. W. POWELL, B.Sc., Ph.D., Physics Department,
The National Physical Laboratory, Teddington, Middlesex *.

[Received October 29, 1938.]

INTRODUCTION.

DURING the last few years it has been shown that the thermal and electrical conductivities of some groups of alloys can be correlated reasonably well by plotting the thermal conductivity, K , against the product of the electrical conductivity, σ , and the absolute temperature T . Some alloys behave abnormally, but in general the points are found to lie within some 12 per cent. of a line given by

$$K = L\sigma T + k,$$

where L and k are constants which vary with the group of alloys studied. When the thermal conductivity is expressed in gram calories per square cm. per second for 1 cm. thickness and 1° C. difference in temperature, and the electrical conductivity is expressed as reciprocal ohms per cm. cube, available data have been found to lead to the values of L and k given in Table I. for the three main groups of alloys to be studied in this way.

TABLE I.

Type of alloy.	Author.	L .	k .
Copper	{ C. S. Smith and E. W. Palmer ⁽¹⁾ .	{ 0.571×10^{-8}	0.018
Aluminium	{ L. W. Kempf, C. S. Smith and C. S. Taylor ⁽²⁾ .		
Iron	{ R. W. Powell ⁽³⁾ .		
		0.502×10^{-8}	0.003
		0.625×10^{-8}	0.006

A few years ago, on behalf of the Metallurgy Department ⁽⁴⁾, thermal and electrical conductivity determinations were carried out on eight magnesium alloys over the temperature range 20° to 250° C. It is now considered appropriate to give some details of the experimental methods,

* Communicated by the Author.

and to endeavour to correlate the results for alloys of this type in a similar manner.

Description of Alloys.

The alloys were available in the form of rods from 28 to 30 cm. in length and about 1.4 cm. in diameter, which had been forged at an elevated temperature. The rods tested bore the identification numbers given in Table II., and were stated to consist of magnesium, of 99.93 per cent. purity, combined with the percentages of additional elements indicated in the table. Approximate values for the densities of these alloys are included.

TABLE II.

Identification no. of magnesium alloy.	Additional elements, per cent.	Density, gm. per ml.
W. 1630	Cerium *, 10.1.	1.86
W. 1635	Nickel, 5.56.	1.84
W. 1567	Manganese, 2.64.	1.77
W. 1641	Calcium, 2.82.	1.74
W. 1648	{ Cerium *, 2.65. Nickel, 5.36.	} 1.87
W. 1662	{ Cerium *, 3.17. Manganese, 1.8.	
N.P.L. P 2	{ Cerium *, 9.0 approx. Cobalt, 3.0 approx. Manganese, 0.5 approx.	} 1.87
W. 1702	{ Cerium *, 2.2. Cobalt, 2.4. Manganese, 1.6.	
		—

Before the present tests were commenced the thermal expansion coefficients of the rods had been determined. This had involved heating each rod to about 250° C.

Measurement of Electrical Conductivity.

The electrical conductivity of each rod was first measured at room-temperature by passing a current of the order of 5 amperes through the rod, and comparing the potential-drop down a central section of the rod with that across a standardized resistance connected in the same simple electrical circuit. Two thermocouples, composed of no. 40 gauge nichrome and eureka wires, pegged into the rod at a distance of about 14 cm. apart, enabled the temperature of the rod to be determined

* The cerium added was in the form of Mischmetal.

and also served as potential leads. The potential readings were taken with the current flowing alternately in positive and negative directions to enable any thermoelectric forces to be eliminated.

In testing the first six alloys mentioned in Table II. the rod was enclosed in a heated fire-clay tube, and electrical conductivity measurements were made up to a temperature of about 250° C. before the thermal conductivity determinations were commenced. For the remaining two alloys, however, the electrical conductivity determinations at temperatures above atmospheric were made at the same time as those of thermal conductivity. This course was adopted after the room-temperature conductivity of some samples had been observed to increase by about 5 per cent. as the result of the prolonged heat treatment to which they were submitted during the thermal conductivity tests.

Values obtained at temperatures of 20°, 50°, 150°, and 250° C. for the electrical conductivity of each rod during heating are set out in Table III. This table also contains values for the reciprocal quantity, the electrical resistivity, which in each case is found to increase approximately as a linear function of the temperature.

Measurement of Thermal Conductivity.

For the purposes of this test an axial hole 10 mm. in diameter was drilled to a depth of about 40 mm. at each end of each rod. One hole was fitted with a heating coil of platinum wire wound on a former of baked steatite, and the other received a water-cooled unit. The latter was in the form of a hollow cylindrical brass vessel, which made good thermal contact with the rod. Water in-flow and out-flow tubes were soldered to the exposed end of this unit, which served as a water-flow calorimeter during the test. Differentially connected thermocouples were included in the water circuit to measure the temperature difference between the in-flowing and out-flowing streams of water. Thermocouples were pegged into the rod at six points along its length.

The specimen was mounted in a vertical position within a metal guard-tube 42 mm. in diameter, the interspace between the rod and guard-tube being packed with heat insulating powder to prevent heat losses from the rod. The base of the guard-tube was also water-cooled, and its top and centre could be heated by heating coils wound as girdles on the tube. The temperature of the guard-tube was also measured by means of thermocouples pegged into the wall.

In carrying out the test a steady temperature distribution was maintained in the rod, lateral leakage of heat being prevented by suitable adjustment of the temperature distribution in the guard-tube.

The temperature gradient in the rod was obtained from readings of the thermocouples pegged into its surface, whilst the quantity of heat

TABLE III.
Electrical Conductivities and Specific Resistances
of Magnesium Alloys.

Identification no. of magnesium alloy.	Temperature (° C.).	Electrical conductivity, σ (ohm. ⁻¹ cm. ⁻¹).	Electrical resistivity, ρ (ohm. cm.).
W. 1630	20	14.9×10^4	6.7×10^{-6}
	50	13.5	7.4
	150	10.3	9.7
	250	8.4	11.9
	20*	15.4	6.5
W. 1635	20	21.3×10^4	4.7×10^{-6}
	50	19.2	5.2
	150	14.3	7.0
	250	11.4	8.7 ₅
	20*	21.3	4.7
W. 1567	20	20.4×10^4	4.9×10^{-6}
	50	18.5	5.4
	150	14.0	7.1 ₅
	250	11.2	8.9
	20*	20.6	4.8 ₅
W. 1648	20	18.5×10^4	5.4×10^{-6}
	50	17.0	5.9
	150	13.0	7.7
	250	10.3	9.7
	20*	19.2	5.2
W. 1641	20	20.6×10^4	$4.8_5 \times 10^{-6}$
	50	18.7	5.3 ₅
	150	14.2	7.0 ₅
	250	11.5	8.7
	20*	—	—
W. 1662	20	16.0×10^4	$6.2_5 \times 10^{-6}$
	50	14.7	6.8
	150	11.8	8.5
	250	9.8	10.2 ₅
	20*	17.1	5.8 ₅
N.P.L. P 2	20	16.4×10^4	6.1×10^{-6}
	50	15.0	6.6 ₅
	150	11.8	8.5
	250	9.6	10.4
	20*	17.4	5.7 ₅
W. 1702	20	18.7×10^4	$5.3_5 \times 10^{-6}$
	50	17.0	5.9
	150	12.8	7.8
	250	10.4	9.6
	20*	19.0	5.2 ₅

* Repeat determination made after completing the thermal conductivity measurements.

producing the gradient was measured both in terms of the electrical input at the heated end and the rate of flow and rise in temperature of the cooling water during its passage through the cooling unit.

In those cases in which the temperature match between the specimen and the guard-tube was not perfect a correction was applied to these two energy measurements to allow for any interchange of heat between the specimen and guard-tube, which occurred between the point at which the energy was measured and the mid-point between the two thermocouples from the readings of which the thermal conductivity was derived. In general the two values obtained in this way for the quantity of heat flowing in the specimen at this point agreed to within 3 per cent. The thermal conductivity was then calculated from the equation

$$K = \frac{Qs}{A(T_1 - T_2)},$$

where Q , expressed in gram calories per second, is the mean of the two values obtained as described above for the energy flowing in the rod, A is the cross-sectional area of the rod in sq. cm., and T_1 and T_2 the temperatures in degrees C. of two points on the rod at a distance apart of s cm.

The experimental results obtained for each alloy over the temperature range of approximately 40° to 270° C. have been plotted separately, and the best smooth curve drawn through each set of points obtained while heating to successively higher temperatures. The extreme deviation of individual points from the curve was rarely more than 3 per cent. Definitely higher values were obtained on cooling for those alloys which also showed an increase in electrical conductivity after being heated to 270° C. for some two or three days.

Table IV. contains values of the thermal conductivity as read from these curves at temperatures of 50° , 150° , and 250° C. A repeat value at 50° C. is included in those cases where the heat treatment associated with the test was observed to cause an appreciable change in conductivity.

Discussion of the Results for Thermal and Electrical Conductivity.

The last column of Table IV. contains values for the Lorenz function of each alloy. The Lorenz function, L , is the ratio of the thermal conductivity to the product of the electrical conductivity and the absolute temperature ($L = K/\sigma T$), and, according to Sommerfeld's theory, this should have a constant value of $0.58_5 \times 10^{-8}$ if electrons are assumed responsible for both thermal and electrical conduction. In the present experiments the values of the Lorenz functions are seen to range from 0.55×10^{-8} to $0.59_5 \times 10^{-8}$, and to have a mean value of $0.57_4 \times 10^{-8}$.

The conductivities of relatively pure samples of magnesium have been

previously examined by several authors over a similar range of temperature. These results are set out in Table V., and the Lorenz function is included in the last column.

TABLE IV.

Thermal Conductivities and Lorenz Functions of Magnesium Alloys.

Identification no. of magnesium alloy.	Temperature (° C.).	Thermal conductivity, K (gm. cal./cm. sec. ° C.).	Lorenz function, L (K/σT).
W. 1630	{ 50 150 250	0.25 0.25 ₅ 0.26	0.57 × 10 ⁻⁸ 0.58 ₅ 0.59
W. 1635	{ 50 150 250	0.34 0.34 0.35	0.55 × 10 ⁻⁸ 0.56 0.58 ₅
W. 1567	{ 50 150 250	0.33 0.34 0.34	0.55 × 10 ⁻⁸ 0.57 ₅ 0.58
W. 1648	{ 50 150 250 50*	0.31 0.31 0.31 0.32	0.56 ₅ × 10 ⁻⁸ 0.56 ₅ 0.57 ₅ 0.57
W. 1641	{ 50 150 250	0.33 ₅ 0.34 0.34	0.55 ₅ × 10 ⁻⁸ 0.56 ₅ 0.56 ₅
W. 1662	{ 50 150 250 50*	0.28 0.29 0.29 ₅ 0.29 ₅	0.59 × 10 ⁻⁸ 0.58 0.58 0.58 ₅
N.P.L. P 2	{ 50 150 250 50*	0.28 0.29 0.30 0.29 ₅	0.58 × 10 ⁻⁸ 0.58 ₅ 0.59 ₅ 0.58
W. 1702	{ 50 150 250	0.31 ₅ 0.31 ₅ 0.31 ₅	0.57 ₅ × 10 ⁻⁸ 0.58 0.58

It will be seen that these results indicate the mean value of the Lorenz function of magnesium to be 0.568×10^{-8} . This value is reasonably

close to the values obtained for the present series of magnesium alloys, and from such limited data the conclusion is reached that the adoption of a value of 0.57×10^{-8} should enable the thermal conductivities of magnesium and its alloys to be derived to within about 5 per cent. from measurements of their electrical conductivities.

Kikuchi ⁽⁹⁾ has also determined the thermal and electrical conductivities of twenty-two magnesium alloys at atmospheric temperature.

TABLE V.

Previously published Results for relatively Pure Magnesium.

Author.	Temperature (° C.).	Electrical conductivity, σ (ohm. ⁻¹ cm. ⁻¹).	Thermal conductivity, K (gm. cal./cm. sec. ° C.).	Lorenz function, $L = K/\sigma T$.
Lorenz ⁽⁵⁾	0	24.47×10^4	0.376	0.563×10^{-8}
„	100	17.50	0.376	0.576
Schofield ⁽⁶⁾ . .	156	13.65	0.328	0.560
„	261	10.73	0.318	0.555
„	326	9.42	0.309	0.548
„	456	7.56	0.314	0.570
Masumoto ⁽⁷⁾ .	29	22.12	0.354	0.530
Mannchen ⁽⁸⁾ .	0	25.58	0.411	0.588
„	100	18.0	0.398	0.593
„	203	13.75	0.389	0.594
Kikuchi ⁽⁹⁾ . .	18	23.1	0.382	0.568

In this series alloys of much lower conductivity were included, which depart considerably from the foregoing conclusion. Thus, the addition of 8.2 per cent. of aluminium gave an alloy for which a Lorenz function of 0.71×10^{-8} was obtained. A correlation of the type mentioned in the introduction becomes necessary when such poorer conductors are included.

In fig. 1 the results for the thermal conductivities of the eight magnesium alloys contained in Table IV. have been plotted against the corresponding electrical conductivities. The points obtained at the three temperatures 50°, 150°, and 250° C. are in each case found to lie about a separate straight line. The equations representing these lines are

$$K = 0.167 \times 10^{-5} \sigma + 0.027 \text{ (at } 50^\circ \text{ C.)},$$

$$K = 0.226 \times 10^{-5} \sigma + 0.023 \text{ (at } 150^\circ \text{ C.)},$$

$$K = 0.275 \times 10^{-5} \sigma + 0.030 \text{ (at } 250^\circ \text{ C.)}.$$

Where T is the absolute temperature these equations reduce to

$K=0.516 \times 10^{-8} \sigma T + 0.027$ (at 50° C.),

$K=0.534 \times 10^{-8} \sigma T + 0.023$ (at 150° C.),

and

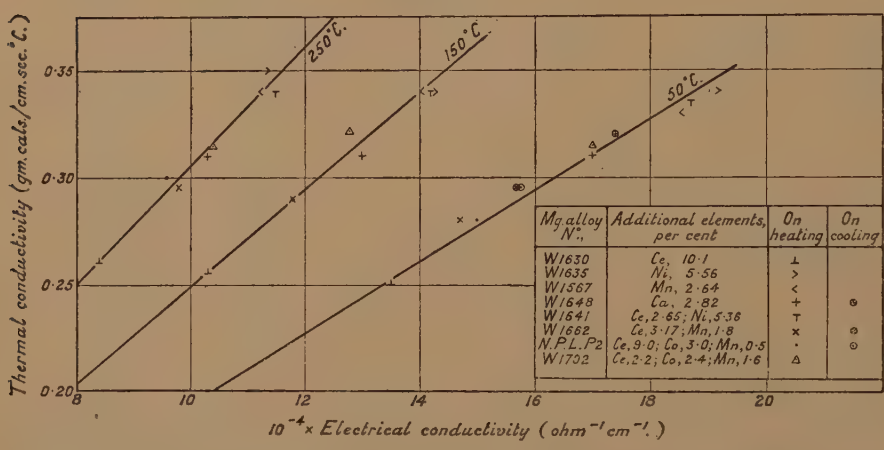
$K=0.527 \times 10^{-8} \sigma T + 0.030$ (at 250° C.).

The equation obtained as the mean of these three equations is

$K=0.526 \times 10^{-8} \sigma T + 0.027.$

Fig. 2 contains a plot of the thermal conductivity against the product of electrical conductivity and absolute temperature, in which existing results for magnesium and its alloys are included. The short dotted line

Fig. 1.



Thermal conductivities at 50°, 150°, and 250° C. against the corresponding electrical conductivities.

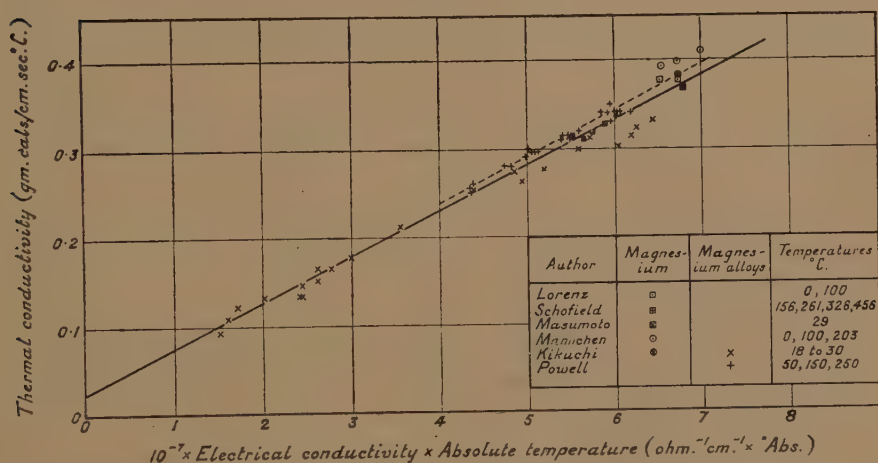
is that represented by the last equation, and it will be seen that all results of the present experiments lie within 3.5 per cent. of this line, whilst those for pure magnesium do not differ from it by more than 5 per cent. It will also be seen that Kikuchi's results for alloys of lower conductivity lie reasonably well about an extension of this line, but that his results for the alloys of higher conductivity are some 10 per cent. low. Values reported by Mannchen ⁽⁸⁾ for a series of magnesium alloys have been omitted from this figure, as they show considerable scatter and differ appreciably from the results of other workers in those instances where comparisons can be made for alloys having similar compositions. This course also receives support from the analysis by Kempf, Smith, and Taylor ⁽²⁾ of the conductivities of aluminium alloys, in which they

remark, "Mannchen's determinations show so much scatter as to suggest, especially in view of the results of other investigators, that some uncontrolled factor exerted considerable influence on his determinations." Through the points plotted in fig. 2 a straight line has been drawn, which is represented by the equation

$$K = 0.516 \times 10^{-8} \sigma T + 0.022.$$

Towards the higher conductivity end this line passes through the mean of Schofield's four determinations, which were obtained by experiments conducted with the utmost care. The specimen studied was stated to be of only 99.6 per cent. purity, so may actually be classed as a rich

Fig. 2.



Thermal conductivity against the product of electrical conductivity and the absolute temperature.

magnesium alloy. The other results for relatively pure magnesium, and all the results of the present investigation for magnesium alloys, lie within about 6 per cent. of this line, and mainly on the high thermal conductivity side. These are balanced by the lower thermal conductivities obtained by Kikuchi for some of the good conducting alloys. Four of his results at this end of the line are more than 6 per cent. low, but they all lie within 10 per cent. of the line. Towards the lower conductivity end Kikuchi's data give points which are evenly distributed about the line and do not depart from it by more than 10 per cent.

It appears that the equation last suggested can be used to determine an approximate value for the thermal conductivity of a magnesium alloy from a measurement of its electrical conductivity, and that the

value so obtained is not likely to differ from the true value by more than 10 per cent., and may be correct to within about 5 per cent.

SUMMARY.

Results are given for the thermal and electrical conductivities of eight magnesium alloys at temperatures of 50°, 150°, and 250° C. It is shown that these results can be represented to within 3·5 per cent. by the equation $K = 0.526 \times 10^{-8} \sigma T + 0.027$, where K is the thermal conductivity expressed in gram calories per square cm. per second for 1 cm. thickness and 1° C. difference in temperature, σ is the electrical conductivity expressed in reciprocal ohms per cm. cube, and T is the absolute temperature. An analysis is then made of existing data for the conductivities of magnesium and magnesium alloys, from which it is concluded that the equation $K = 0.516 \times 10^{-8} \sigma T + 0.022$ can be used to give an approximate value for the thermal conductivities of such metals.

ACKNOWLEDGMENT.

The author desires to express his thanks to Mr. M. J. Hickman, who carried out most of the experimental work.

References.

- (1) C. S. Smith and E. W. Palmer, Amer. Inst. Mining and Metal, Engineers. Tech. Publn. no. 648 (1935).
- (2) L. W. Kempf, C. S. Smith, and C. S. Taylor, Amer. Inst. Mining and Metal, Engineers Trans. cxxiv., Inst. Metals Division, p. 287 (1937).
- (3) R. W. Powell (in course of publication).
- (4) J. L. Haughton and W. E. Prytherch, 'Magnesium and its Alloys' (London, H.M.S.O. 1937), p. 62.
- (5) L. Lorenz, Wied. Ann. xiii. pp. 422 & 582 (1881).
- (6) F. H. Schofield, Proc. Roy. Soc. A, cvii. p. 206 (1925).
- (7) H. Masumoto, Sci. Rep. Tôhoku Univ. xiii. p. 229 (1925).
- (8) W. Mannchen, Zeits. f. Metallkunde, xxiii. p. 193 (1931).
- (9) R. Kikuchi, Sci. Rep. Tôhoku Univ. xxi. p. 585 (1932).

he, in agreement with other investigators, observed that friction, as distinct from seizing, of polished, was less than that of etched bismuth. W. B. Hardy and J. K. Hardy⁽⁵⁾ laid stress on the wear of surfaces during sliding, but W. B. Hardy and Doubleday⁽⁶⁾ quite correctly pointed out that Amontons' law "holds so long as the surfaces, or rather the whole system, solid surfaces and lubricant, remain unchanged," so that "cases in which the traction alters the texture of the solid" are outside the range of Amontons' law, and thus of ordinary friction. It may be mentioned here that Coulomb⁽¹⁾ discontinued his measurements whenever scratching of the surfaces became apparent.

Hardy and Doubleday raised "an unanswerable objection to Coulomb's theory," namely that it did not explain the difference in lubricating properties between acids and alcohols on one hand and paraffins on the other. But we might ascribe this difference as due to the high spreading capacity of the polar substances which fill up the depressions in the solid surfaces and do not form droplets like the benzene droplets described by Hardy and Doubleday⁽⁹⁾.

Hardy was apparently convinced that as the friction took place between surfaces, surface forces must play a predominant role. But it might be observed, as Landsberg⁽¹⁰⁾ has pointed out, that strong surface forces are usually neutralized by adsorption of air and are not involved in ordinary friction.

Whilst these objections raised against Coulomb's theory do not appear to be very convincing, Coulomb's criterion requires reconsideration when the difference between the real area A_0 and the geometrical or apparent area A of contact is taken into account. The total adhesion between two bodies is, of course, proportional to A_0 , although only A can be measured with precision. If A_0 is in fact proportional to the load equations (1) and (2) become identical, and Amontons' empirical law would be consistent with both mechanisms. A proportionality between A_0 and L has indeed been postulated by both Tomlinson⁽²⁾ and by Bowden and Tabor⁽³⁾ *.

In the present work another criterion is adopted. If the frictional force f_1 between a slider of weight W_1 and a plate is first determined, and then an additional weight W_2 be added to the slider, the new frictional force f_2 would, if Amontons' law holds, be

$$f_2 = f_1 \frac{W_1 + W_2}{W_1}.$$

If the additional weight W_2 be retained for a short period on the slider and subsequently removed, it is evident that the new frictional force

* Cf. also Holm⁽¹²⁾.

f_3 should return to the original value f_1 if it is due to Coulomb's second cause including any elastic deformation of the surface. If, on the other hand, the frictional force is due to adhesion, this will not be affected by *reduction* of the load from $W_1 + W_2$ to W_1 , and consequently f_3 should be equal to f_2 and not to f_1 . This equality of f_3 and f_2 should hold if the friction is due to true adhesion or to what may be described as a welding⁽⁴⁾ or as a plastic flow⁽³⁾, for the removal of the extra load W_2 does not introduce any force which could act in such a manner as to separate the bodies in contact.

2. Experimental.

The frictional force was determined in the usual way from the equation $f = W \sin \alpha$ by measuring the angle required to produce sliding, W being the weight of the sliding body (*i. e.*, slider or slider + additional weight). Precautions were taken to avoid casual overloading on addition, or disturbance on removal of the extra load. For if welding or plastic flow occurs, a slider deposited with appreciable velocity would show a higher friction than one placed carefully on the surface. To obviate this possibility the sliding body, after being placed on the surface, was pushed sideways to a small extent, thus breaking the "welded points," if any, formed at the moment of fall. To ensure that the removal of the extra weight could be effected without disturbing the boundary between the slider and support, the additional weight was suspended on the arm of a hand-balance, so that it could be lifted by a weight W_2 put on the other scale. Since the coefficient of friction μ is of the order 0.1 to 0.5, movement of the weight is only possible if the horizontal component of the force involved in raising the extra weight is larger than 0.1 ($W_1 + W_2$). This horizontal component was made small by fixing the thread*, on which the extra weight was suspended, over the centre of gravity of the latter; and the sum ($W_1 + W_2$) was made sufficiently large as compared with W_2 by taking W_2 approximately equal to the weight of slider W_1 . Too small an additional weight would entail too small a difference between f_1 and f_2 , and therefore between the two theoretical values of f_3 .

All the surfaces, excepting those of paraffin wax, were cleaned with soap and water, followed by distilled water and petroleum ether†. No attempt has been made to prevent "lubrication" of the surface by air and impurities contained in it, as our purpose was to examine the mechanism of friction under ordinary conditions.

The sliders and the additional weights were thick and approximately square plates. Their masses are given below.

* By experiment, it was shown that any adventitious torsion of the thread was insufficient to displace the slider.

† Free from grease, when tested by the surface film technique.

Ground brass : slider 3.80 g., additional weight 3.79 g.

Plate glass : slider 9.38 g., additional weight 9.65 g.

Polished and chromium-plated brass : slider 27.50 g., additional weight 27.99 g.

Two paraffin wax (congealing temperature 60°) sliders 0.175 and 0.57 g., with the additional loads 0.20 and 0.68 g.

The supporting plates included :

Ground brass (grinding was carried out till the reflection of a scale could readily be observed in the surface).

Glass (glass side of a photographic plate).

Chromium-plated polished brass.

Tin-plate.

Paraffin wax (a molten and solidified surface).

	f_1 .	f_2 .	f_3 .	μ_1 .	μ_2 .
Brass on tin-plate	1.24	2.48	1.28	0.34	0.33
Brass on wax	1.07	2.08	1.07	0.295	0.27
Brass on brass	0.965	1.94	0.99	0.262	0.264
Brass on glass	0.92	1.67	0.87	0.25	0.225
Brass on chromium	0.593	1.17	0.581	0.158	0.156
Glass on tin-plate	2.77	5.56	2.87	0.31	0.29
Glass on glass	2.30	4.50	2.26	0.25	0.23
Glass on brass	1.66	3.54	1.75	0.18	0.18
Glass on chromium	1.57	3.18	1.65	0.17	0.16
Chromium on chromium	3.04	6.12	3.08	0.115	0.112
Chromium on brass	2.89	5.71	2.97	0.106	0.103

The measurements of the friction between each pair of surfaces included a survey of the effect of time of contact. The angle of sliding was determined after the slider (with or without additional weight) was in contact with the supporting surface for 2, 10, 30, 100, 100, 30, 10, and 2 seconds, respectively. The combination of pairs investigated revealed no appreciable time effect, and all subsequent readings were made after a contact of from five to ten seconds.

The frictional forces were measured in the order $f_1 \rightarrow f_2 \rightarrow f_3 \rightarrow f_3 \rightarrow f_2 \rightarrow f_1$; after four or five such sets of determinations had been made, the surfaces were cleaned and then four or five sets of readings were again taken.

The average values thus obtained are given in the Table. The f and μ values were calculated from the equations $f_1 = W_1 \sin \alpha_1$, $f_2 = (W_2 + W_1) \sin \alpha_2$, $f_3 = W_1 \sin \alpha_3$, $\mu_1 = \tan \alpha_1$, $\mu_2 = \tan \alpha_2$, α_1 being the sliding angle for slider alone, α_2 that for slider plus additional weight, and α_3 that after removal of the additional weight. f values are given in grams (not in dynes).

The table shows that for hard surfaces f_3 is within the limits of the experimental error equal to f_1 , thus eliminating cohesion, "welding," and plastic flow as factors in the friction. A comparison of the coefficients of friction μ_1 and μ_2 shows that for the same surfaces Amontons' law is approximately valid.

For surfaces deformed as easily as those of paraffin wax even the small pressures used (about 500 dynes per sq. cm.) were sufficient to cause some adhesion; this fact was reflected simultaneously both in a deviation from Amontons' law and in the existence of an after-effect. Two examples for sliding of wax on wax may be given. For the small slider $f_3:f_1=1.45$, $\mu_1=0.35$, and $\mu_2=0.30$, and for the larger slider $f_3:f_1=1.47$, $\mu_1=0.25$, and $\mu_2=0.22$ were obtained. The small slider on tin-plate gave $f_3:f_1=1.23$, $\mu_1=0.31$, and $\mu_2=0.28$.

These experimental results may be expressed in the statement that, as long as Amontons' law is valid, the friction is not due to cohesion, so that we are free to attribute it to elevation of the slider over the roughnesses or asperities in the surface. It is interesting to note that this theory gives us not only the form of the law of friction ($f=\mu L$), but also the absolute value of the constant*. In presence of air or another "lubricant" all the usual values of μ lie within the limits of 0.1 and 0.5, corresponding to the angles of sliding of 6° and 27° respectively. On Coulomb's view we must conclude that even surfaces which are highly polished must contain irregularities forming angles of about 6° with the visible surface, and that when these angles become greater than about 27° the surfaces cease to be "smooth" and are not suitable for measurement of friction. From the point of view of the cohesion theory, any value of μ might be anticipated.

Summary.

If sliding friction is due to cohesion, "welding," or plastic flow, the value of the coefficient of friction must depend on the previous overloading of the surfaces in contact. As this after-effect cannot be detected for hard surfaces, the friction between hard surfaces cannot be due to the above-mentioned effects. On the other hand, Coulomb's theory ascribing friction to a lifting of the slider over irregularities finds a semi-quantitative confirmation in the absolute value of the coefficients of friction.

References.

- (1) Coulomb, "Théorie des machines simples." Paris, 1809. (Reprint from *Mémoires des Savants Etrangers*, t. (x).)
- (2) Tomlinson, Phil. Mag. (7) vii. p. 905 (1929).

* It can also be applied to the friction between drops of liquid and solid surfaces (11).

- (3) F. P. Bowden and D. Tabor, *Proc. Roy. Soc. A*, clxix. p. 391 (1939).
- (4) N. K. Adam, 'The Physics and Chemistry of Surfaces,' Oxford, p. 219 (1938).
- (5) W. B. Hardy and J. K. Hardy, *Phil. Mag.* (6) xxxviii. p. 32 (1919).
- (6) W. B. Hardy and I. Doubleday, *Proc. Roy. Soc. A*, c. p. 550 (1922).
- (7) W. B. Hardy, *Phil. Mag.* (6), xl. p. 201 (1920).
- (8) Technical Bulletin issued by E. G. Acheson, Ltd., No. 220, p. 2 (1934).
- (9) W. B. Hardy and I. Doubleday, *Proc. Roy. Soc. A*, c. pp. 565-6 (1922).
- (10) K. E. Landsberg, *Ann. d. Physik and Chemie*, cxxi. p. 283 (1864).
- (11) J. J. Bikerman, *Proc. Roy. Soc. A*, clxx. p. 130 (1939).
- (12) R. Holm, *Wiss. Veroff. Siemens Werke*, xvii. p. 38 (1938).

LXV. *The Interpretation of the Michelson-Morley and
Kennedy-Thorndike Experiments.*

By HERBERT DINGLE, A.R.C.S., D.Sc., Professor of Natural
Philosophy, Imperial College of Science and Technology *.

[Received April 4, 1939.]

Introductory.

A REMARKABLE series of papers by Dr. H. E. Ives has recently appeared †, in which the significance of the Michelson-Morley and Kennedy-Thorndike experiments on the absolute motion of the earth is carefully analysed and the conclusion is reached that the relativity explanation is unsatisfactory, the facts requiring the existence of a fixed ether in which instruments are modified by motion. Although this subject has been widely discussed, there is still much divergence of opinion—and, as even recent literature shows, some confusion of thought—concerning the proper interpretation of the experiments. Objections to the relativity theory are still maintained on the ground either that its conceptions are metaphysical and therefore inadmissible, or that it does not agree with experiment; and, further, many physicists appear to hold the view that the hypotheses of relativity and of the Fitzgerald-Lorentz contraction are simply different ways of saying the same thing. In these circumstances, a searching analysis such as that conducted by Dr Ives is most desirable, and it is to be hoped that the opportunity which it affords of reaching the truth on the matter will not be allowed to pass without an attempt to reach general agreement.

The purpose of this paper is twofold: first, to express the ideas of the special theory of relativity in a rather unusual way, so as to emphasize their simplicity, their fundamentally physical character, and their absolute distinction from the idea of the Fitzgerald-Lorentz contraction; and, secondly, to show that Dr Ives's objections to the theory, and his interpretation of the experiments, are inadmissible.

* Communicated by the Author.

† (I.) J. O. S. A. xxvii. p. 177 (1937); (II.) xxvii. p. 263 (1937); (III.) xxvii. p. 305 (1937); (IV.) xxvii. p. 310 (1937); (V.) xxvii. p. 389 (1937); (VI.) xxviii. p. 296 (1938). When necessary, these papers will be referred to in the text by the corresponding Roman numerals.

SER. 7, VOL. 27, NO. 185—JUNE 1939.

2 Z

The Experiments and the Alternative Interpretations.

The experiments are too well known to need detailed description, but a very brief statement of their essential principle is necessary. In each of them a beam of monochromatic light is divided into two parts which travel to and fro along directions at right angles to one another and produce interference fringes on reunion. In the Michelson-Morley experiment the paths of the beams were made equal, as measured by rods at rest with respect to the apparatus and the Earth, and in the Kennedy-Thorndike experiment they were made unequal. According to classical ideas there should be a shift of the fringes as the apparatus is rotated, depending on the speed and direction of the Earth's motion through the ether; but no shift was observed, apart from relatively small experimental errors and possible displacements of the kind suggested by Prof. Dayton Miller, the reality of which is irrelevant to the present controversy.

Two of the proposed explanations have survived. According to the first, every material body, when moving through the ether with velocity v , is shortened in the direction of motion in the proportion $(1 - v^2/c^2)^{\frac{1}{2}} : 1$, where c is the velocity of light. This shortening (the "Fitzgerald-Lorentz contraction") would explain the Michelson-Morley experiment, but not the Kennedy-Thorndike experiment. Ives calculates, however (I., p. 179), that both experiments are explained if we assume that bodies are altered in length in the proportions $[(1 - v^2/c^2)^{\frac{1}{2}}]^n + 1 : 1$ in the direction of motion and $[(1 - v^2/c^2)^{\frac{1}{2}}]^n : 1$ in perpendicular directions; and that moving clocks are altered in frequency in the ratio $[(1 - v^2/c^2)^{\frac{1}{2}}]^{1-n} : 1$. A very skilful experiment performed by Ives and Stilwell* shows that the only admissible value of n is zero. This means that the alteration of lengths with motion reduces to the Fitzgerald-Lorentz contraction and, in addition, that moving clocks must be supposed to change in frequency in the proportion $(1 - v^2/c^2)^{\frac{1}{2}} : 1$.

The second explanation (the special theory of relativity) is simply that absolute motion is meaningless. The experiments then need no apology, being obviously in accord with the hypothesis, and the only purpose in discussing them is to discover what mistake led us to assume that they would reveal absolute motion. Such discussion, supplemented by the result of the Ives-Stilwell experiment, shows that the mistake has been the omission of a factor $(1 - v^2/c^2)^{\frac{1}{2}}$ from all physical relations in which length is concerned either explicitly or implicitly; that the quantity l , in fact, expressing the result of measuring a length by a scale at rest with respect to the body measured, never occurs alone in laws of nature, but always as one term of the expression $l(1 - v^2/c^2)^{\frac{1}{2}}$, where v is the

* J. O. S. A., xxviii. p. 215 (1938).

velocity of the body, reckoned with respect to whatever arbitrary position of rest may be chosen, in the direction of the length l .

That this is not, as may at first appear, a highly artificial procedure but, on the contrary, one which in its general form is perfectly natural and might have been expected, will be shown presently ; but first it is desirable to express the contrast between the two explanations as sharply as possible. According to the first, v has a definite, absolute value, whatever false standard of rest we may choose for convenience in particular problems. It is the velocity with respect to a fixed ether, and measuring rods and clocks undergo a definite physical change when their value of v is altered. We have not been able to detect this change directly because some compensating effect, peculiar to each experiment made for the purpose, has always occurred, but nevertheless, it is real, and is as much a physical property of each individual moving body as are the observable shape and other characteristics of the body. According to the relativity explanation, however, v is entirely conventional. There is no meaning in absolute motion, so that velocity can be conceived only with respect to an arbitrary standard. If we choose a standard of rest such that the body in question has no motion, then $v=0$, and the expression $l(1-v^2/c^2)^{\frac{1}{2}}$ reduces to l . If, instead, we choose another standard of rest, then v is no longer zero and the full expression must be used. But nothing whatever has happened to the body through this change ; the process is purely mental. On the contraction hypothesis we change the length of a body by changing its state of motion : on the relativity hypothesis we change the form of our equations by changing our mind *. The difference between the two hypotheses may therefore be summed up in this way : *on the contraction hypothesis, a measuring rod or a clock undergoes an intrinsic physical change when its velocity is changed ; on the relativity hypothesis, change of velocity of the instrument is a conceptual process involving no intrinsic physical change.*

The Character of the Relativity Explanation.

The significance of the hypothesis that absolute motion is meaningless becomes clearer when we compare it with the parallel case of relativity of position. It is generally agreed that there is no absolute position in space, and that the position we assign to any body must be relative to a selected zero position. If, now, we try to determine the position of a particle (for simplicity we consider one dimension only) merely by laying a graduated scale through it and taking the reading, x , against which it

* It must be remembered, of course, that we cannot change the *relative* velocities of bodies in this way ; it is only the absolute motion of each single body that is in question.

lies, we shall get no significant result. We must observe also the reading, a , corresponding to the chosen zero point, and then the position of the particle will be represented by $x-a$. The quantity a here is precisely analogous to the quantity $(1-v^2/c^2)^{\frac{1}{2}}$ in relation to the relativity of motion. We can make it what we please, and the x of the particle will change accordingly without the occurrence of any physical change. It is, in fact, because a is arbitrary that it *must* appear in all expressions of the position of the particle in order to save the truly physical results from arbitrariness. Any attempt to determine its absolute value must always fail because it has none; in every equation expressing a physical relation between particles it cancels out and so evades determination. There is obviously nothing metaphysical in this; it is non-physical only in the sense in which the principles of arithmetic are non-physical. It is true that in the problem of position we can see at once how the quantity a should appear, whereas in the problem of velocity it is not obvious why v should have to be introduced in the form which we find to be necessary. This, however, is doubtless due to our lack of insight, resulting from the traditional adoption of a false point of view from which absolute velocity appears to have meaning. A sufficiently clear mind should be able to see at once, without having to deduce the fact laboriously from experiment, that all measurements of length should be multiplied by the factor $(1-v^2/c^2)^{\frac{1}{2}}$; but our present deficiencies in this respect do not affect the obviousness of the general principle that if there is no absolute zero of velocity, but for purposes of convenience we choose an arbitrary one, *some* reference to our choice must appear in our equations in order to preserve them from being themselves arbitrary.

If the legitimacy of the relativity hypothesis as a physical explanation is granted, its great simplicity as compared with the alternative gives it a strong claim to adoption, provided that it does not conflict with experiment. This proviso, which Ives maintains that it does not satisfy, will be considered in the next paragraph. Anticipating a favourable conclusion, it is worth while here to point out how much simpler it is than the hypothesis proposed by Ives. It consists of the single assumption that absolute motion is meaningless. Even if it be conceded that the multiplication of all measured lengths by $(1-v^2/c^2)^{\frac{1}{2}}$ is an additional assumption (which, as already stated, we need not admit inasmuch as some such procedure is a necessary corollary), the whole postulatory basis of the theory thus expressed is still far simpler than that of the hypothesis of instrument modification. The additional term introduced with time, mass, and other measurements is deducible from that introduced with space measurements, because the definitions of those measurements involve measurement of space: no additional postulate is necessary, and certainly no metaphysical assumption about the "inner nature" of the

things measured. Take time, for example. The basis of the method adopted for measuring time is Newton's first law of motion. Having chosen a standard unit of space measurement, we take a body moving freely under the influence of no external forces *, and define equal times as those during which it covers equal amounts of space. It follows that since the space covered must always be multiplied by a factor which has the effect of reducing it as v is increased, the moving body constituting the "hand," so to speak, of the clock will, as v is increased, effectively cover a smaller space in a given duration; in other words, the measure of that duration will be reduced by the same factor, which gives the same result as if the clock ran slow. The effect on other physical measurements follows in the same way.

In contrast to this simple hypothesis that absolute motion is meaningless (augmented, if thought necessary, by the "correction" to moving lengths), consider the complication of the Ives hypothesis, which adds to the Fitzgerald-Lorentz contraction the assumption that moving clocks run slow in a stated proportion. The postulates, as stated by Ives ((II.) p. 269), are as follows:—

"I. The velocity of light in the luminiferous ether, when measured by light signals between two points at rest in the ether, by clocks and rods unaffected by transport, is a constant c .

"II. The lengths of material rods (in the direction of motion) and the frequencies of material clocks, are reduced by the factor $(1-v^2/c^2)^{\frac{1}{2}}$, where v is the velocity of the rod or clock, and c the velocity of light, measured as in I."

These postulates are set out as though there were only two of them, but actually there are an indefinitely large number. We may grant that the Fitzgerald-Lorentz contraction is a single assumption, since all matter is believed to have the same fundamental constitution; and we may even, to avoid irrelevant argument, take it without prejudice as a deduction from the electrical theory of matter and therefore not an assumption at all. But this by no means holds for the postulate about "the frequencies of material clocks." "Material clocks" is a vague term. The rotating Earth is a "clock," and so are a vibrating quartz

* The ambiguity of the phrase, "under the influence of no external forces," which is of importance in the *general* theory of relativity, has no significance here, for the only essential point for our present purpose is that equal times are defined in terms of equal spaces; it does not matter what moving body traverses those spaces. All actual clocks are attempts to realize the Newtonian ideal, and if we have reason to suspect imperfect realization—as, for instance, in the case of the rotating Earth on account of tidal friction—we make "corrections" to the clock readings before accepting the results.

crystal, a swinging pendulum, a radiating atom, an hour glass, a burning candle, dripping water, and innumerable other devices. There is no standard clock as there is a standard metre, for the unit of time, as we have seen, is an ideal conception realized in a greater or less degree by a large variety of very different instruments. You cannot set an ideal conception moving and examine its rate. You can do that only with a material clock, and if you make an assumption about the behaviour in such circumstances of one kind of material clock, it certainly does not follow that another must behave so as to give you the same measure of time. If the frequency of a swinging pendulum changes in a certain proportion when moving, you cannot deduce that the frequency of an hour glass will change in the same proportion: that must be a separate assumption. It follows that Ives is making a set of quite independent assumptions about an indefinitely large number of "material clocks," and none of them is relevant because all readings are in any case "corrected" to reach a conceptual ideal.

It does not seem possible, by any hypothesis which involves actual physical changes in measuring instruments, to meet the facts of experiment without multiplying *ad hoc* assumptions to the point of absurdity. The Fitzgerald-Lorentz contraction will explain the Michelson-Morley experiment, but not the Kennedy-Thorndike experiment, and the addendum which Ives proposes turns out to be such a large group of independent hypotheses as to discredit itself on grounds of mental economy. We shall now see that it is further discredited by the fact that it does not even fulfil its purpose of explaining the Kennedy-Thorndike experiment.

The Alternative Hypotheses in Relation to Experiment.

The principle of the Kennedy-Thorndike experiment is extremely simple, the only measurements involved being those of two lengths and of the position of a set of interference fringes. It is therefore immediately obvious that no hypothesis about physical modifications of material clocks can have anything to do with the matter, for no material clock was used; and, further, since the fringes were undisplaced, no assumptions about time measurement are involved in the interpretation. It only remains, therefore, to locate the defect in Ives's reasoning.

The method by which he shows that a null result would be expected on his hypothesis is as follows ((I.) p. 180). He converts the path-difference of the beams (corrected for contraction) into a time-difference by dividing it by c , and multiplies the result by his clock-correction factor to produce "the *measured* time interval between arrival of reflected pulses at the origin." The result is independent of v , and it is concluded that therefore the experiment would be expected to give a negative

result. But such a calculation is quite irrelevant to the experiment. There is no measured time interval, because the times of arrival of the beams are not measured ; all that is measured are the lengths of the arms and the zero displacement of the fringes. What Ives has done is to assume that *if* the times of travel of the beams had been measured by a clock, and *if* the velocity of light obeyed his first postulate, and *if* the clock had been modified by motion in a manner postulated *ad hoc*, the difference of times would have appeared to be independent of the velocity of the apparatus. There is nothing here but assumptions ; the *fact* is simply that the fringes were undisplaced. It may or may not be valid to deduce a modification of clocks from the occurrence of unexpected observations, but it can scarcely be valid to deduce it from the incompatibility of gratuitous assumptions*.

So far as experiment is concerned, then, the position appears to be as follows. Omitting all attempts to detect absolute motion other than the two experiments here discussed, the Michelson-Morley experiment may be accounted for either by the theory of relativity (that absolute motion is meaningless) or by the hypothesis of the Fitzgerald-Lorentz contraction (that motion through the ether produces a shortening of material bodies) ; but the Kennedy-Thorndike experiment can be accounted for only by the theory of relativity. Before, however, we can conclude that the theory of relativity offers the only satisfactory account of facts at present known, we must consider certain objections which Ives has advanced.

1. *Absolute Simultaneity*.—It is well known that the denial of meaning to absolute motion implies the denial of meaning to absolute simultaneity of spatially separated events. On this point Ives writes as follows ((II.) p. 270) :—

“ The objection is commonly stated in terms of the idea of simultaneity at separated points. If this cannot be determined, it is argued that it has ‘no meaning.’ Now the distinction which must be made here is between *nonexistence* and *indeterminacy*. . . . The *existence* of true simultaneity is, however, easily established by the following observation :—

“ Consider three clocks *A*, *B* and *B'*, all having the same rates when together, and let *B* and *B'* be moved to a distant point on the common moving platform. Send a series of equidistant pulses from *B* to *A*, and let *A* be set to be synchronous with the pulses it receives. Then let the pulses sent from *A* be used to set *B'*. We will then have *B* and *B'* going

* The assumption that the times taken by light beams are actually measured by material clocks is repeated in the discussion (VI.) of the Sagnac experiment, which therefore does not need independent consideration here.

at the same rate, but with a difference of phase. Let us now set up a number of other clocks at B , B' all of the same rate, but at phases between B and B' . Now by taking a sufficiency of these clocks, *one will be* beating simultaneously with A , irrespective of the fact that we cannot determine which it is, within the limits set by B and B' , which give a measure of the indeterminacy."

Now this again is to make the matter appear more subtle than it actually is. What relativity denies is that the simultaneity in question has meaning: what Ives has shown is that, *if* it has meaning, then it can be realized in nature. But that is not the point. We may equally well prove the existence of an absolute zero of position: put a particle at every point in space, and by taking a sufficiency of particles, one will be at the absolute zero, irrespective of the fact that we cannot determine which it is. Similarly, we may prove that there is a magic letter. Write down the whole alphabet, and then one letter *must* be the magic letter, although we cannot tell which. The relativity contention can be met only by challenging its *criterion of significance*, which is that unless a conception can be related directly or indirectly to experience, it has no significance*.

2. *The Aberration of Light*.—In the same paper ((II.) p. 271), Ives maintains that the fact of aberration of light from the stars is evidence for the existence of the luminiferous ether, *i. e.*, of a medium with respect to which absolute motion is significant. But again the matter is much simpler than it is represented. The aberration of light is simply a name given to the phenomenon of a periodic change in the relative positions of the Earth and a star, and it reveals only the relative motion of the Earth and the star. Its interpretation in terms of a postulated energy pattern in a postulated ether is perfectly legitimate as a means of describing the phenomenon, but it cannot possibly give evidence for the *existence* of such postulates, because they have to be assumed before the argument begins. The fact of aberration could give evidence for the existence of Ives's ether only if it could yield information about the motion of a body with respect to that ether, and this it cannot do. The further argument that it is enormously more probable that the Earth moves through the energy pattern than that the energy pattern moves past the Earth is a still further complication of the essential simplicity of the question. It becomes relevant only if it is supposed that one of these alternatives is true and the other false; and this, which relativity denies, Ives does not discuss.

3. *The Michelson-Morley Experiment and the Ether*.—Another instance of needless complication of the question is contained in the statement

* For a discussion of the wider implications of this criterion, see Proc. Roy. Inst. xxx. p. 68 (1937), or an abbreviated account in 'Nature,' cxli. p. 21 (1938).

((II.) p. 271) that the Michelson-Morley experiment proves nothing with regard to the invariance of light signals with the velocity of the apparatus, unless it is assumed that the Earth has absolute motion. "Calling upon the Michelson-Morley experiment as proof of the invariance of light signal phenomena therefore carries with it the acceptance of the luminiferous ether here assumed." But relativity does not call upon the Michelson-Morley experiment for such proof: it regards the experiment simply as one among many failures to detect the undetectable. The factor $(1-v^2/c^2)^{\frac{1}{2}}$ makes a null result inevitable, whatever value v may have, including zero. If you assume that a light signal is an energy pattern travelling through a fixed ether, then relativity requires you to assume also that light signal phenomena are invariant, but it leaves you perfectly free to make that assumption or not. It cannot, therefore, carry with it "the acceptance of the luminiferous ether here assumed."

4. *The Clock Paradox*.—Ives describes "the physical and logical absurdity of the 'clock paradox'" as "a consequence of a sweeping and unqualified application of the hypothesis that relative motion of matter is the only operative factor" ((III.) pp. 307-9). The "paradox" is described as follows:—

"Given two similar clocks A and B ; let B be moved to a distant point at the velocity V with respect to A , and back to A . It will, because of its motion, go at a slower rate, a function of V^2/C^2 , and will on its return be *slow with respect to A*. But if relative motion only is of significance, A is likewise moving at the velocity V with respect to B , and hence will, when the clocks are together again, be *slow with respect to B*. Obviously two clocks, side by side, cannot each be slow with respect to the other. There is here a logical inconsistency which demands an examination of the premises."

Now here once more the extreme simplicity of the matter from the point of view of relativity has been involved in extraneous subtleties. Relativity says simply that absolute motion is meaningless: hence the relative motion of A and B may equally well be described as a motion of A with respect to B or a motion of B with respect to A ; that is all. If, then, the clocks show different times at the end of the experiment, they must have been treated differently; and obviously they have. Clocks do not naturally move apart and come together again: in order to make them do so, some force must be applied to one and not to the other, and the difference in reading must therefore be due to the difference in physical treatment. Relativity does not predict what clocks will do when constrained to move unnaturally: it simply says that the *motions* involved may be described with equal validity with respect to any assumed standard of rest.

There is therefore no "paradox"; there is simply a *problem* concerning the effect on clock readings of the application of certain forces. This problem has been discussed by Tolman *, who uses the *general* theory of relativity to calculate the difference of readings and to show that the same difference is obtained whichever clock is assumed to be moving, provided that in each case it is the same clock which has its natural motion interfered with.

Since this paper is almost entirely critical of Dr. Ives's work, I should not like to conclude without expressing appreciation of the care with which he has analysed the experiments under discussion, and particularly of the brilliant additional experimental evidence with which he has been associated and which will be of permanent value. By far the greater part of his work is supplementary to, rather than contradictory of, the relativity hypothesis, and my difference from him consists mainly in the fact that I do not think he has realized this. The principle of relativity is of precisely the same character as the second law of thermodynamics—a negative statement based on all the relevant evidence available. It can never be proved that absolute motion is meaningless, any more than it can be proved that heat cannot move spontaneously against a temperature gradient; whereas either hypothesis might be disproved by a single observation. Those who accept the relativity hypothesis, however, regard experiments designed to detect absolute motion in the same light as those who accept the second law of thermodynamics regard perpetual motion machines. They are ready to consider any experimental evidence that absolute motion has meaning, but they no longer consider that a detailed examination of the reasons why particular experiments must fail is a profitable occupation. Nevertheless, when an examination as thorough as that conducted by Dr. Ives is made, it is impossible not to feel an additional satisfaction that from whatever point of view the Michelson-Morley and Kennedy-Thorndike experiments are regarded, they must fail to reveal absolute motion.

* 'Relativity, Thermodynamics, and Cosmology,' p. 194.

LXVI. *Some Formulæ for the Associated Legendre Functions of the First Kind.*

By T. M. MACROBERT, Professor of Mathematics,
University of Glasgow *.

[Received January 30, 1939.]

§ 1. *Introductory.*

IN § 2 of this paper an integral involving an associated Legendre Function is evaluated by means of a known formula for a Generalized Hypergeometric Function. In § 3 some further applications are made of the method of deducing Legendre Function formulæ from Bessel Function formulæ, developed in a previous paper (Phil. Mag. xxi. pp. 697-703 (1936)). The references below to "Hobson" are to Hobson's treatise on 'Spherical and Ellipsoidal Harmonics.' Similarly, "G.M.M." refers to 'Bessel Functions' by Gray, Mathews, and MacRobert.

§ 2. *Evaluation of a Definite Integral involving an Associated Legendre Function of the First Kind.*

The formula to be proved is

$$\int_0^\pi (\sin \theta)^{l-1} T_n^{-m}(\cos \theta) d\theta = \frac{\pi \Gamma\left(\frac{l+m}{2}\right) \Gamma\left(\frac{l-m}{2}\right)}{2^m \Gamma\left(\frac{l+n+1}{2}\right) \Gamma\left(\frac{l-n}{2}\right) \Gamma\left(\frac{m+n}{2}+1\right) \Gamma\left(\frac{m-n+1}{2}\right)}, \quad (1)$$

where $R(l+m) > 0$ and

$$T_n^{-m}(x) = \frac{(1-x^2)^{\frac{1}{2}m}}{2^m \Gamma(m+1)} F\left(\begin{matrix} m-n, m+n+1 \\ m+1 \end{matrix}; \frac{1-x}{2}\right). \quad (2)$$

If in the integral we put

$$\cos \theta = 1 - 2\mu,$$

it becomes

$$\begin{aligned} & \frac{2^{l-1}}{\Gamma(m+1)} \int_0^1 \{\mu(1-\mu)\}^{\frac{1}{2}l+\frac{1}{2}m-1} F\left(\begin{matrix} m-n, m+n+1 \\ m+1 \end{matrix}; \mu\right) d\mu \\ &= \frac{2^{l-1}}{\Gamma(m+1)} B\left(\frac{l+m}{2}, \frac{l+m}{2}\right) {}_3F_2\left(\begin{matrix} m-n, m+n+1, \frac{1}{2}l+\frac{1}{2}m \\ m+1, l+m \end{matrix}; 1\right). \end{aligned}$$

* Communicated by the Author.

On applying the formula

$${}_3F_2 \left\{ \begin{matrix} a, b, c; 1 \\ \frac{1}{2}(a+b+1), 2c \end{matrix} \right\} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}+c)\Gamma(\frac{1}{2}+\frac{1}{2}a+\frac{1}{2}b)\Gamma(\frac{1}{2}-\frac{1}{2}a-\frac{1}{2}b+c)}{\Gamma(\frac{1}{2}+\frac{1}{2}a)\Gamma(\frac{1}{2}+\frac{1}{2}b)\Gamma(\frac{1}{2}-\frac{1}{2}a+c)\Gamma(\frac{1}{2}-\frac{1}{2}b+c)}. \quad (3)$$

and the formula

$$\Gamma(\frac{1}{2})\Gamma(l+m) = 2^{l+m-1} \Gamma\left(\frac{l+m}{2}\right) \Gamma\left(\frac{l+m+1}{2}\right), \quad (4)$$

formula (1) is obtained. Formula (3) was given by Whipple (Proc. Lond. Math. Soc. xxiii. p. 113 (1923)). It is a generalization of a particular case given earlier by G. N. Watson (see Bailey, 'Generalized Hypergeometric Series,' page 16).

§ 3. Proofs of some Integral Representations of Associated Legendre Functions of the First Kind.

The formula

$$\sqrt{\left(\frac{\pi}{2}\right)} \Gamma(m+n) \Gamma(m-n) (z^2-1)^{\frac{1}{2}(\frac{1}{2}-m)} P_{n-\frac{1}{2}}^{\frac{1}{2}-m}(z) = \int_0^\infty e^{-\lambda z} K_n(\lambda) \lambda^{m-1} d\lambda \quad (5)$$

is valid for $R(z) > -1$, $R(m \pm n) > 0$. It can be established by expressing $K_n(\lambda)$ in terms of $I_n(\lambda)$ and $I_{-n}(\lambda)$ and integrating term by term. The formula

$$P_n^{-m}(z) = \frac{\sin(n-m)\pi}{\pi \cos n\pi} \frac{\Gamma(n-m+1)}{\Gamma(n+m+1)} \{Q_n^m(z) - Q_{-n-1}^m(z)\} \quad (6)$$

is required in the proof.

Now in (5) substitute the formula (G.M.M., p. 51)

$$K_n(\lambda) = \int_0^\infty e^{-\lambda \cosh t} \cosh(nt) dt, \quad (7)$$

where $R(\lambda) > 0$. Then, on changing the order of integration, we find that

$$\sqrt{\left(\frac{\pi}{2}\right)} \Gamma(m+n) \Gamma(m-n) (z^2-1)^{\frac{1}{2}(\frac{1}{2}-m)} P_{n-\frac{1}{2}}^{\frac{1}{2}-m}(z) = \Gamma(m) \int_0^\infty \frac{\cosh(nt) dt}{(z + \cosh t)^m}, \quad (8)$$

where $R(z) > -1$, $R(m \pm n) > 0$ [Hobson, p. 262 (121)].

Similarly, by employing the formula (G.M.M., p. 50)

$$K_n(\lambda) = \frac{\sqrt{\pi}}{\Gamma(n+\frac{1}{2})} \left(\frac{\lambda}{2}\right)^n \int_0^\infty e^{-\lambda \cosh \phi} (\sinh \phi)^{2n} d\phi, \quad (9)$$

where $R(\lambda) > 0$, $R(n+\frac{1}{2}) > 0$, the formula

$$2^{n-\frac{1}{2}} \Gamma(m-n) \Gamma(n+\frac{1}{2}) (z^2-1)^{\frac{1}{2}(\frac{1}{2}-m)} P_{n-\frac{1}{2}}^{\frac{1}{2}-m}(z) = \int_0^\infty \frac{(\sinh \phi)^{2n} d\phi}{(z + \cosh \phi)^{n+m}}, \quad (10)$$

where $R(z) > -1$, $R(n+\frac{1}{2}) > 0$, $R(m-n) > 0$, is obtained.

The application of Whipple's Transformation to (8) and (10) leads to corresponding formulæ for $Q_n^m(z)$ [Hobson, p. 259 (117), p. 256 (111)].

Note.—In formulæ (5), (8), and (10)

$$(z^2 - 1)^{\frac{1}{2}(\frac{1}{2} - m)} P_{n - \frac{1}{2}}^{\frac{1}{2} - m}(z)$$

should be replaced by

$$(1 - z^2)^{\frac{1}{2}(\frac{1}{2} - m)} T_{n - \frac{1}{2}}^{\frac{1}{2} - m}(z)$$

when $|z| < 1$. The two functions are identical, and are uniform and continuous at $z = 1$.

LXVII. *Note on a Problem in Heat Conduction.*

By HAROLD W. WOOLLEY *.

[Received January 2, 1939.]

IT may be worth while to point out an error in the paper "Certain Vibration Problems solved by means of an Analogous Problem in Heat Conduction," by George Green, *Phil. Mag.* xxii. p. 1079 (Dec. 1936). The particular problem in heat conduction considered is the determination of the temperature at any point and at any instant in a conducting rod of uniform cross-section, one end of which ($x=0$) is maintained at zero temperature and the other end of which ($x=a$) is in perfect thermal contact with a heat reservoir of finite heat capacity. The system is initially at zero temperature throughout. After a certain instant heat is supplied at a constant rate, Q , to the heat reservoir. Heat loss from the surface of the rod is disregarded.

We have the following definitions :

K = thermal conductivity of the rod.

A = area of cross-section of the rod.

a = length of the rod.

κ = thermal diffusivity of the rod.

Ms = heat capacity of the reservoir.

The problem is treated first by a method of wave trains and next by the ordinary method based on the form of the solution obtained by the first method.

In this second method the temperature is given by

$$\theta = \theta_1 + \frac{Qx}{KA},$$

where it is assumed that

$$\theta_1 = \sum_{n=1}^{\infty} B_n e^{-\kappa \lambda_n^2 t} \sin \lambda_n x,$$

λ denoting λ_n , the positive roots of the equation

$$\tan \lambda a = \frac{KA}{Ms\kappa\lambda}.$$

* Communicated by Lyman J. Briggs.

As stated by Green, the two methods of solution of the problem do not appear to give the same result, except in case

$$\sum_{n=1}^{\infty} B_n \sin \lambda_n a = 0.$$

It is apparent, however, that

$$\sum_{n=1}^{\infty} B_n \sin \lambda_n a$$

is the value of θ_1 at time $t=0$, and at $x=a$. Accordingly

$$\sum_{n=1}^{\infty} B_n \sin \lambda_n a$$

has the value

$$-\frac{Qa}{KA}$$

and cannot be zero.

The error is in the part on the wave-train method in the discussion preparatory to equation (27) on page 1084, in which the strength of instantaneous sources q includes the heat in the rod but not that in the finite heat capacity Ms at $x=a$. The inclusion of the contribution due to the finite heat capacity Ms will alter equation (27), which is incorrect, and make it consistent with equation (33). The latter equation is undoubtedly correct, whereas Green stated that "there is no doubt that the result given in (27) is correct."

This problem was of interest to the writer in a study of transient temperatures occurring in the use of the guarded hot plate for measuring thermal conductivity. A solution was obtained by a third method using contour integration and based on a paper by Carslaw*.

This method is limited to the case where the initial temperature is uniform, a restriction which is not necessary in the two methods previously mentioned, but may be of interest as leading to further confirmation of equation (33) and to an explicit solution of the stated problem.

Consider the expression

$$\theta = \frac{Q}{i\pi} \int \frac{\sin \lambda x e^{-\kappa \lambda^2 t} d\lambda}{(KA \cos \lambda a - Ms\kappa \lambda \sin \lambda a) \lambda^2},$$

where the integration is along the standard contour above the real axis. By closing the curve of integration in different ways it may be shown that this expression satisfies the boundary conditions and may be reduced to the form

$$\theta = \frac{Qx}{KA} - 2Q \sum_{n=1}^{\infty} \frac{e^{-\kappa \lambda_n^2 t} \sin \lambda_n x}{\lambda_n^2 \sin \lambda_n a} \frac{KA}{a\{(KA)^2 + (Ms\kappa \lambda_n)^2\} + Ms\kappa \lambda_n a},$$

* "Bromwich's Method of Solving Problems in the Conduction of Heat," H. S. Carslaw, Phil. Mag. vol. xxxix. p. 603 (May 1920).

where λ denotes λ_n , the positive roots of

$$\tan \lambda a = \frac{KA}{Ms\kappa\lambda}.$$

The term $\frac{Qx}{KA}$ is obtained automatically from the residue for $\lambda=0$.

The result may also be written as

$$\theta = \frac{Qx}{KA} - 2Q \sum_{n=1}^{\infty} \frac{e^{-\kappa\lambda^2 t} \sin \lambda x \sin \lambda a}{\lambda^2 [Ms\kappa \sin^2 \lambda a + aKA]},$$

or as

$$\theta = \frac{Qx}{KA} - \frac{2Q}{KA} \sum_{n=1}^{\infty} \frac{e^{-\kappa\lambda^2 t} \sin \lambda x \sin \lambda a}{\lambda^2} \frac{(KA)^2 + (Ms\kappa\lambda)^2}{a\{(KA)^2 + (Ms\kappa\lambda)^2\} + Ms\kappa KA}.$$

LXVIII. *Calculation of Triode Constants.*

By J. H. FREMLIN, M.A., Ph.D., A.Inst.P.*

[Received March 17, 1939.]

SUMMARY.

A new treatment of the equivalent diode is proposed, from which formulæ for anode current and mutual conductance in plane or cylindrical triodes are obtained in terms of the penetration factor, the inter-electrode distances, and the voltages applied to grid and anode. In the plane case the values found differ from those given by the most widely used expressions by a factor

$$\left[\frac{1+D}{1+D \left(\frac{l_a}{l_g} \right)^{4/3}} \right]^{3/2},$$

which may in some circumstances be considerably less than 1.

In the cylindrical case a similar but more complex factor is obtained. The expressions obtained for the plane case are shown experimentally to be more nearly correct than previous expressions.

These formulæ can only be used in cases for which the penetration factor can be calculated. Beginning with Maxwell's expression for the potential distribution due to a charged grid of fine wires it is shown how this distribution can be calculated for the case of a grid close to the cathode (Appendix D), and hence the value of penetration factor can be calculated for zero current at these close spacings.

It is shown experimentally that the value of penetration factor thus obtained may be used with some accuracy up to considerable current densities. Calculated and experimental values of mutual conductance at close spacings are compared, and it is shown that if allowance is made for emission velocity very close agreement can be obtained down to values of grid-cathode clearance of only a few per cent. of the grid pitch. A definite maximum value of mutual conductance for constant current density is shown to occur at a cathode-grid clearance of the order of half the grid pitch.

* Communicated by Prof. E. V. Appleton, F.R.S.

INTRODUCTION.

IT is of fundamental importance in the design and in the development of thermionic valves to be able to determine in advance what the characteristics of such valves will be. In the case of a diode it is already possible to do this to a very considerable accuracy when the electrodes are either parallel planes or concentric cylinders. Where there are one or more grids between the anode and the cathode, however, the position is less satisfactory. Schottky, Miller ⁽¹⁾, King ⁽²⁾, and others have derived expressions for the amplification factor μ of an infinite plane triode and of a cylindrical triode of infinite length in which the ratio $\frac{\text{wire diameter}}{\text{grid pitch}}$,

d/a , is assumed small (less than 0.1). Vogdes and Elder ⁽³⁾ have obtained a formula available for larger values of d/a up to about one-third, and Ollendorf ⁽⁴⁾ gives a method of calculation said to be valid for all values of d/a . The theoretical values of μ agree very well with experiment so long as the distance of the grid from the cathode l_g is greater than about one grid pitch.

Calculations of anode current per unit area i and mutual conductance per unit area g_m are not usually as successful, though a large number of formulæ have been proposed both on theoretical and on empirical grounds. The work described in this paper was undertaken in the attempt to make more accurate calculations possible for triodes. It was especially concerned with the case in which the grid is very close to the cathode, for which relatively little experimental work has been published. Before an expression could be developed for this case it was necessary to investigate the expressions for current in the simpler conditions where the wires of the grid are thin and where the cathode-grid distance is appreciably more than one grid pitch.

In Part I. theoretical formulæ are developed for i and g_m when the cathode-grid distance is greater than the grid pitch, and the apparatus used for testing these is described.

In Part II. it is shown how these formulæ may be used when the grid is very close to the cathode so long as the reduction of penetration factor is taken into account *.

Definitions of the symbols used and a table of references are given at the end in Appendices A and B.

* Throughout this paper the quantity "penetration factor," sometimes described as "Durchgriff," is in general used in preference to its reciprocal, the amplification factor. The use of the term "amplification factor" seems rather inapt in many cases, particularly in the consideration of electrostatic problems, and is therefore confined to cases in which its physical meaning is clear.

PART I.

Theory.

The expressions most frequently used ⁽⁵⁾ for anode current and mutual conductance are (omitting a small contact-potential term) :

$$i = \frac{2.34 \times 10^{-6} (V_g^2 + DV_a)^{3/2}}{l_g^2 (1+D)^{3/2}} \text{ amps. per unit area, } \quad \text{. . . (I.)}$$

where $V_a V_g$ are the anode and grid potentials, D is the penetration factor (reciprocal of the amplification factor), and l_g is the distance of the plane of the grid from that of the cathode, and

$$g_m = \frac{\partial i}{\partial V_g} = \frac{3.51 \times 10^{-6} \sqrt{V_g + DV_a}}{l_g^2 (1+D)^{3/2}} \text{ amps. per volt per unit area} \quad \text{(II.)}$$

for the plane case, and

$$i = \frac{1.47 \times 10^{-5} (V_g + DV_a)^{3/2}}{r_g \beta_{cg}^2 (1+D)^{3/2}} \text{ amps. per unit length, } \quad \text{(III.)}$$

$$g_m = \frac{2.20 \times 10^{-5} \sqrt{V_g + DV_a}}{r_g \beta_{cg}^2 (1+D)^{3/2}} \text{ amps. per volt per unit length} \quad \text{(IV.)}$$

for the cylindrical case ⁽⁵⁾. Here r_g is the radius of the grid cylinder and β_{cg}^2 is a function of r_g/r_c (where r_c is the cathode radius) given by Langmuir ⁽⁷⁾.

These formulæ were not calculated directly for a triode, but indirectly as follows. Child ⁽⁶⁾ and Langmuir ⁽⁷⁾ showed how the space-charge limited current could be calculated through any diode, the electrodes of which were infinite parallel planes or infinite concentric cylinders. Then, if it could be proved that the current through a given triode was exactly that which would pass through a particular diode whose dimensions and anode potential were known, this current would be known. The particular diode is called the "equivalent diode." Its electrode parameters naturally depend upon those of the triode to which it is equivalent, and if the conception is to be of value they must be calculable in terms of these.

The formulæ (I.) and (II.) above were calculated on the simple assumption, without detailed analysis, that the diode with cathode-anode distance l_D given by

$$l_D = l_g \quad \text{. (V. a)}$$

and with anode voltage V_D given by

$$V_D = V_g + DV_a \quad \text{. (V. b)}$$

was nearly equivalent to the triode. The derivation of the correction

factor $\frac{1}{(1+D)^{3/2}}$ is given by (among others) Chaffee ⁽⁸⁾; the reasoning seems at times to be somewhat obscure. Formulæ (III.) and (IV.) were found similarly, assuming $r_D = r_g$. Other authors have made different but still arbitrary assumptions. For example, Miller ⁽¹⁾ assumed that a diode giving the same cathode field as the triode in the absence of space-charge, and having its anode in the same position as the anode of the triode, would give the same current. His formula is as follows:

$$i = \frac{2.34 \times 10^{-6} (V_g + D V_a)^{3/2}}{\sqrt{l_a} (l_g + D l_a)^{3/2}} \quad \text{amperes per unit area.}$$

Owing to the assumption that

$$l_D = l_a$$

this underestimates the current considerably, and is not much used.

Benham ⁽¹³⁾ has recently put forward reasons why a formula similar to Miller's, but putting l_D instead of l_a , where

$$l_D = l_g + D l_a,$$

should be accepted. Unlike Miller's formula, this ensures that the denominator tends correctly to l_g^2 as D becomes very small.

For large values of D (0.1 or above, say, for ordinary valve dimensions) these formulæ, particularly Benham's, give more nearly the current which is determined experimentally than does formula (I.) given above. It does not, however, seem to have been realized that it is possible to calculate the current through a triode without direct reference to a particular equivalent diode. This can be done as follows. Consider first the plane case.

If a current of density i is flowing between plane electrodes under space-charge limited conditions, the potential distribution will be given by

$$V^{3/2} = k i l^2, \quad \dots \dots \dots \text{(VI.)}$$

where $k = \frac{1}{2.34 \times 10^{-6}}$ if i is measured in amps. per sq. cm. and V is

measured in volts. Here the cathode is taken as origin of V and of l , and the effects of initial velocity of the electrons are neglected. Equation (VI.) then merely expresses the Child-Langmuir law. The distribution is shown in fig. 1. Suppose that the anode and grid of the triode are introduced into the electron stream at such potentials that the original distribution is unaltered, *i. e.*, so that

$$\left. \begin{aligned} V_a &= k^{2/3} i^{2/3} l_a^{4/3}, \\ V_g &= k^{2/3} i^{2/3} l_g^{4/3}. \end{aligned} \right\} \dots \dots \dots \text{(VII.)}$$

We have now a triode with electrode potentials which are known in terms of the total space current which is passing. It is known that the field at the cathode of a plane triode is given by

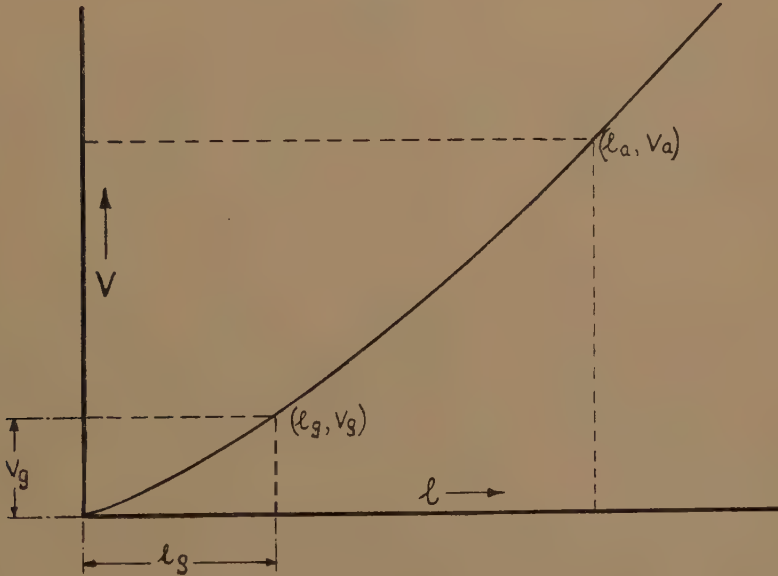
$$\frac{\partial V}{\partial l} = \frac{V_g + DV_a}{l_g + Dl_a} \dots \dots \dots \text{(VIII.)}$$

in the absence of space-charge ⁽⁹⁾. When a space-charge-limited current flows the cathode field is reduced to zero, and it is known that for any electrode system the current required to do this is proportional to the three-halves power of the cathode field. Hence we know that

$$i = K(V_g + DV_a)^{3/2}, \dots \dots \dots \text{(IX.)}$$

where K depends on the dimensions of the valve.

Fig. 1.



The curve represents the potential distribution in a plane triode when the charge on the grid is zero and a space charge limited current i is flowing; the potential V at any distance l from the cathode is then given by $V^{3/2} = ki \cdot l^2$.

Substituting from (VII.) for V_g, V_a , we have

$$i = K(k^{2/3}i^{2/3}l_g^{4/3} + Dk^{2/3}i^{2/3}l_a^{4/3})^{3/2};$$

$$\therefore K = \frac{1}{k(l_g^{4/3} + Dl_a^{4/3})^{3/2}};$$

$$\therefore K = \frac{2.34 \times 10^{-6}}{[l_g^{4/3} + Dl_a^{4/3}]^{3/2}} \text{ in practical units. } \dots \dots \dots \text{(X.)}$$

Hence
$$i = \frac{2.34 \times 10^{-6} (V_g + DV_a)^{3/2}}{[l_g^{4/3} + D l_a^{4/3}]^{3/2}} \text{ amps. per unit area, . . . (XI.)}$$

whence we have

$$g_m = \frac{3.50 \times 10^{-6} \sqrt{V_g + DV_a}}{[l_g^{4/3} + D l_a^{4/3}]^{3/2}} \text{ amps. per volt per unit area. (XII. a)}$$

From (XI.) and (XII. a) we have also

$$g_m = \frac{2.64 \times 10^{-4} \cdot i^{1/3}}{[l_g^{4/3} + D l_a^{4/3}]} \text{ amps. per volt per unit area. . . (XII. b)}$$

These differ from the usual formulæ used only in the replacement of the term $[1 + D]^{3/2}$ by the term

$$\left[1 + D \left(\frac{l_a}{l_g} \right)^{4/3} \right]^{3/2}.$$

It will be noticed that the equation for current has been derived without explicit reference to the equivalent diode. The form of the equations show that the triode is equivalent to a particular diode, however, whose dimensions and anode potential (from equations (VIII.) and (XI.)) are given by

$$l_D = \frac{[l_g^{4/3} + D l_a^{4/3}]^3}{[l_g + D l_a]^3}; \quad V_D = \frac{[l_g^{4/3} + D l_a^{4/3}]^3}{[l_g + D l_a]^4} [V_g + DV_a]. \quad \text{(XIII.)}$$

The difference between this value of l_D and the values hitherto assumed expresses the difference between equation (XI.) and previous formulæ which have been given for current. Diodes giving the same cathode fields in the absence of current, but having different electrode spacings, pass different currents, and the importance of this fact has not been sufficiently realized.

No mention has been made so far either of the fact that part of the current will be stopped by the grid wires when these are positive or of the possible variation of the penetration factor. The effect of the former will always be very small, as if the wires are relatively thin the current will be little affected, while if they are thick the space-charge between grid and anode will anyway be largely shielded from the cathode.

If the penetration factor varies very rapidly the formulæ (XII. a) and (XII. b) may be in error, owing to the existence of appreciable terms involving

$$\frac{\partial D}{\partial V_g} \quad \text{and} \quad \frac{\partial D}{\partial V_a}.$$

The non-uniformity of space-charge owing to focusing by the grid wires when the grid is negative will not affect the field of the space-charge

at the cathode so long as the grid is far enough away for its own field to be uniform.

A similar analysis has been carried out to determine the radius and potential of the anode of the equivalent diode for the case of a cylindrical triode. Given that D can be calculated, the results apply equally well to the squirrel-cage or helical grid in the absence of "inselbildung"*. Large grid supports, however, cause a definite "beam" effect which makes many circular section valves obey the formulæ for planes much more closely than they do those for cylinders. With these provisos then we have

$$i_i = \frac{14.7 \times 10^{-6} (V_g + DV_a)^{3/2}}{D(r_{ai}\beta_{ca}^2)^{2/3} + (r_{gi}\beta_{cg}^2)^{2/3}]^{3/2}} \text{ amps. per unit length, . (XIII.)}$$

$$g_m = \frac{22.0 \times 10^{-6} \sqrt{V_g + DV_a}}{[(r_{gi}\beta_{cg}^2)^{2/3} + D(r_{ai}\beta_{ca}^2)^{2/3}]^{3/2}} \text{ amps. per volt per unit length. (XIV.)}$$

The effect of initial velocities has been neglected throughout this discussion. In the case of the plane diode Langmuir⁽¹⁰⁾ has calculated rigorously the potential distribution during flow of current assuming a Maxwellian distribution of emission velocity among the electrons. It would in principle be possible to calculate the equivalent diode on this basis in a manner similar to the above. It is not possible in practice, however, to obtain separate expressions for l_D and V_D , the spacing and voltage of the equivalent diode, though it can be seen that l_D would in general vary with current.

It is possible to obtain a first approximation to the true values by measuring the distances and voltages from the potential minimum rather than from the cathode itself when the distance l_m of this from the cathode can be calculated. It is not difficult to find l_m if the cathode temperature is known, for currents a small fraction of the saturation current, using the formulæ developed by Langmuir. For large currents, approaching the saturation value, initial velocity effects may be neglected. In Appendix C is estimated the correction required to allow for emission velocities to this degree of approximation.

Experiment.

An apparatus has been built by means of which the grid and anode of a plane triode may be moved about at will while the current is flowing. It was desired that the position of the electrodes with respect to one another should be very accurately adjustable over a range of 2 to 3 centimetres. To do this a short but accurate optical bench was built of steel. This was fixed rigidly to a circular steel base, and covered by

* The term "inselbildung" is here taken to mean the state of affairs in which emission from the cathode surface is appreciably non-uniform.

a bell-jar which was sealed to the base with Apiezon Q. A pumping tube, a tube leading to an ionization manometer, and a tube carrying the necessary insulated electrical leads were soldered into the base, and the whole system could be evacuated by means of an oil-diffusion pump backed by a Cenco Hivac. In figs. 2 and 3 are shown diagrams of the carriages and frames which were used to support anode and grid. In each case to the small steel carriage A sliding upon the optical bench was bolted a sheet of steatite to which the frame M was fixed. The plane of each electrode was adjustable by means of the three screws S, against which the electrode itself was held by a molybdenum spring. The two carriages

Fig. 2.

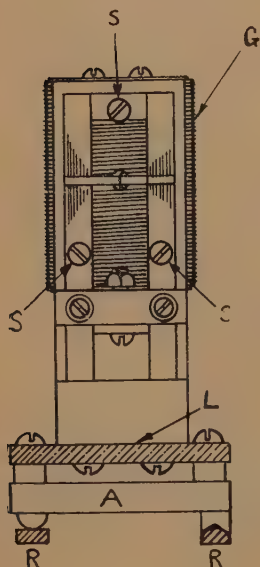


Fig. 3.

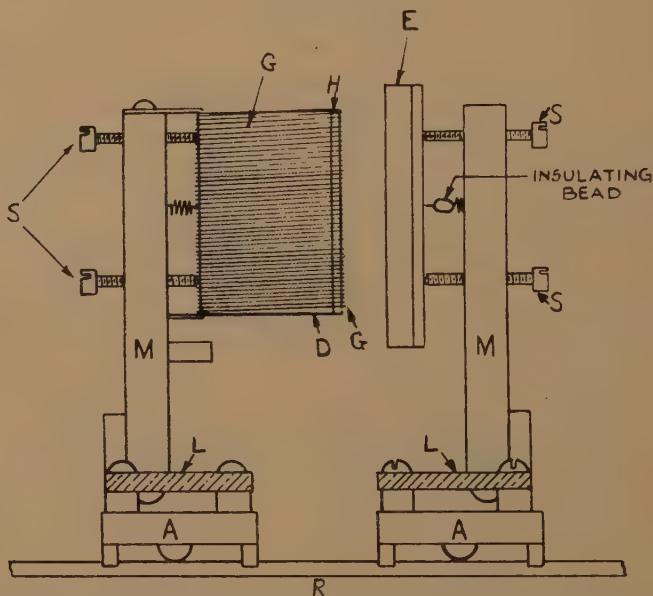


Diagram of the method of mounting the grid and anode in the experimental apparatus. AA, steel carriages; D, steel frame of grid; E, nickel sheet anode; G, molybdenum laterals of grid; H, screwed rod, ensuring constant grid pitch; LL, lavite plates; MM, steel supporting frames; RR, parallel guide rails; SS, adjusting screws against which the electrodes are held by springs.

were moved by means of screwed rods (not shown) operated by steel ground joints lubricated with Apiezon L grease, and the positions of the electrodes were determined by graduations on the moving parts of the ground joints. As the screw used was of exactly 1 mm. pitch, one complete rotation represented 1 mm. movement of the carriage. It was possible to set the position of either electrode repeatedly to an accuracy of 0.01 mm. In fig. 3 the cathode support is omitted for the sake of

clarity ; this was bolted rigidly to the steel base, perpendicular to the steel rails on which the carriages ran. An indirectly heated coated cathode C was used, mounted on steatite insulators attached to a steel supporting frame shown in section at B in fig. 4. The cathode itself was of unconventional type, as it was necessary to ensure absolute flatness. The working part consisted of a solid piece of nickel, 30 mm. by 8 mm. by 5 mm. thick, ground accurately flat on one face. The heater was contained in a shallow channel milled in the back, and was held in by two or three layers of thin nickel sheet, very lightly attached, to reduce radiation losses. As is indicated by the section of the whole electrode system shown in fig. 4, one face only of the cathode C was coated and used.

The cathode C and the frame B (from which it was insulated) were completely surrounded by the light steel frame D carrying the grid wires G. In order that the grid wires should be truly coplanar they were wound upon this frame, of which the edges were very carefully ground flat and parallel. To ensure even pitch two light screwed rods H were welded near the edges of D, and the wires of the grid were wound into the threads at suitable intervals. This avoided the small irregularities always produced by welding. The grid could be adjusted parallel to the cathode extremely well, so that measurements could be made down to a clearing distance of 0.002 cm. between the cathode coating and the inner side of the grid wires, although the working length of the grid was 3 cm.

In order to investigate the variation of current and mutual conductance in a planar triode without disturbing edge effects the anode was made of nickel sheet in guard ring form. A small rectangular section F (fig. 4) was insulated from the main body of the anode E by a thin coating of insulating paste, the whole being baked at red heat *in vacuo* before use. This guard ring structure, though absolutely necessary, was not quite so flat as were the cathode and grid. The position of the centre section F, 1.550 cm. long by 0.371 cm. broad, was known to 0.004 cm., and the anode at no place departed from flatness by more than this amount.

An ionization gauge was used to measure the pressure. Between it and the apparatus was a small trap which could be surrounded by liquid air. This made it easy to distinguish between air and condensable vapour, and made the finding of leaks a practically painless process. During the experimental work the pressure was maintained at between 5×10^{-7} and 3×10^{-6} mm. ; the apparatus could be degassed enough for this by running for about a day after the initial activation of the cathode.

Tests were made to see whether the anode structure was properly carrying out its function of eliminating edge effects. If the edge effects were eliminated changes in them must also have been eliminated. They could readily be changed by altering the potential V_B with respect to

the cathode of the frame B round the cathode (fig. 4) (B was usually maintained at cathode potential). A grid potential of zero and an anode potential of 200 volts were maintained, the grid being in position midway between anode and cathode. A change from -15 to $+15$ volts of V_B produced no detectable change in the current to the centre section of the

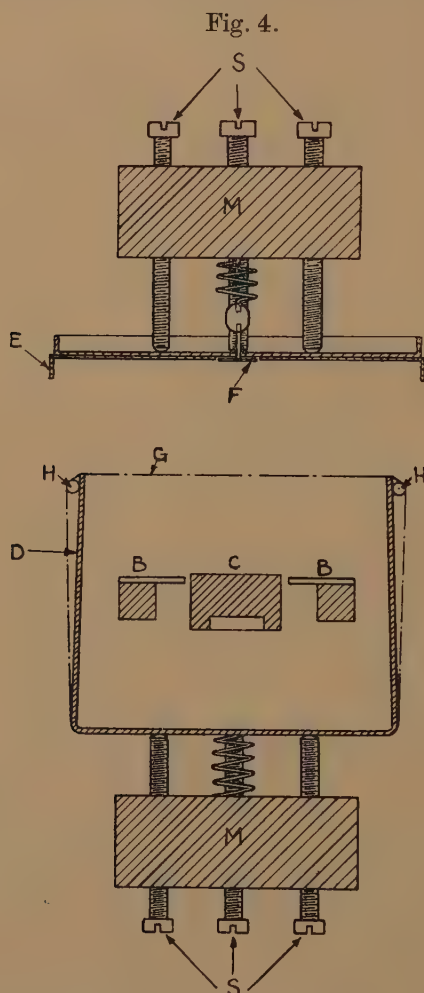


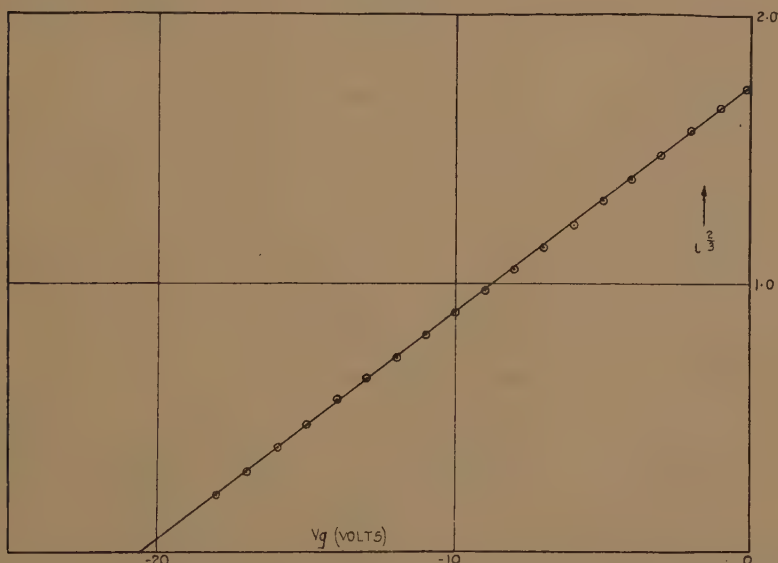
Diagram to show the electrode construction.

BB, shield round cathode; C, indirectly heated nickel cathode; D, steel frame of grid; E, outer "guard" part of the anode; F, centre section of the anode, area 0.575 sq. cm. insulated from E by heater coating; G, molybdenum grid laterals; HH, screwed rods fixing the grid pitch; MM, supporting frames; SS, adjusting screws.

anode, I_F , though the current to the outer part of the anode, I_E , was doubled. Change of V_B from -30 volts to $+30$ volts gave a change of I_F from 1.00 to 1.04 ma., while I_E changed from 5 ma. to 15 ma.

It seems clear from this that the guard ring was effective. The value of I_F was very sensitive, however, owing to secondary emission, to any small potential difference between E and F (fig. 4). A change of 1 volt sometimes made a difference of more than 10 per cent. in I_F . Care was taken therefore during all experimental work to ensure that no potential difference was set up; it was hoped that contact potential differences would be small, as E and F were made from the same sample of nickel and were treated similarly throughout.

Fig. 5.



This shows the variation of anode current density i with grid potential V_g , indicating that the three-halves power law is closely obeyed. Anode potential was 200 volts and the dimensions were as follows: $-l_g = 0.170$ cm.; $l_a = 0.773$ cm.; $a = 0.175$ cm.; $d = 0.0089$ cm.

Figs. 5-9 show some results of the experiments. First, tests were made of amplification factor to see whether the Miller-Schottky formula *

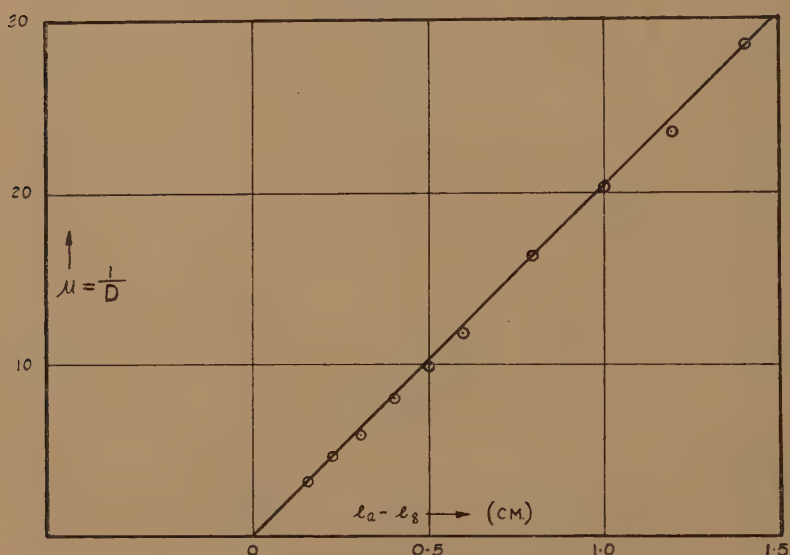
* See ref. (1). For plane electrodes this is
$$\mu = \frac{2\pi(l_a - l_g)}{a \log_e \frac{1}{2 \sin \frac{\pi d}{2a}}}$$
, where

$(l_a - l_g)$ is the distance between grid and anode planes, a is the grid pitch, and d is the grid wire diameter.

could be trusted for the values of d/a , with which it was desired to work. Owing to the double anode construction it was difficult to measure the amplification factor dynamically, and the most accurate method was found to be extrapolation of a line obtained by plotting $I_F^{2/3}$ against V_g to zero current. This method also gave confirmation of the fact that current obeyed the three-halves-power law (see fig. 5).

Fig 6 shows that good agreement between calculated and observed amplification factor was obtained. This being so, the relations obtained for current and mutual conductance could fairly be tested.

Fig. 6.



Variation of penetration factor D and amplification factor μ with grid-anode distance $(l_a - l_g)$.

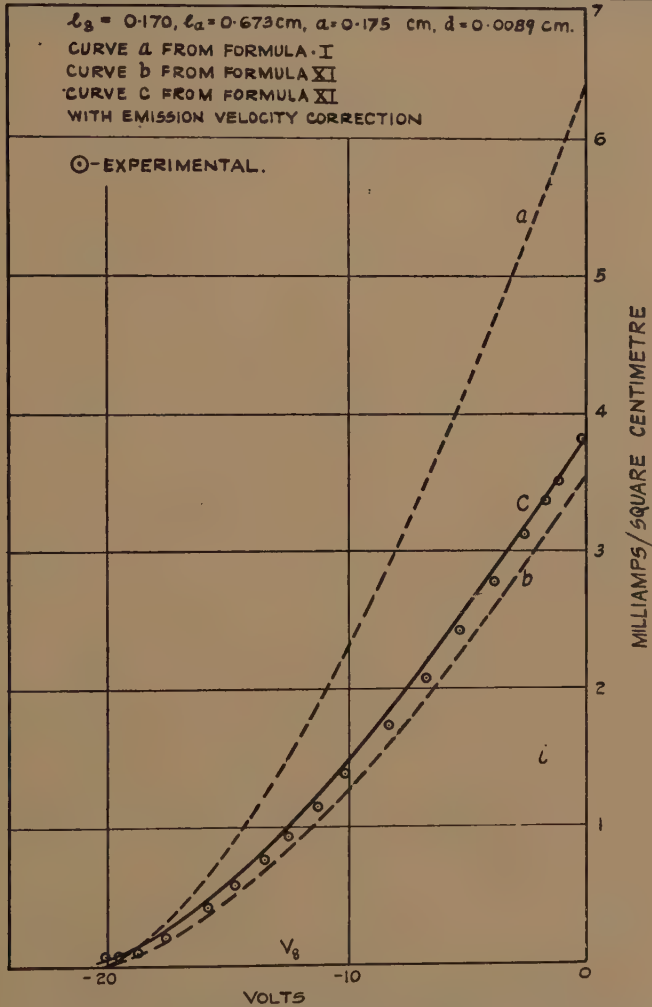
Full line=theoretical. \odot =experimental points.
 $l_g=0.170$ cm., $a=0.175$ cm., $d=0.0089$ cm.

In fig. 7 are shown actual measured values of anode current together with curves calculated from equation (I.) (curve a) and equation (XI.) (curve b). It can be seen that curve (b) is considerably more accurate than (a), but that nevertheless equation (XI.) appreciably underestimates the current. This is very largely accounted for if we take into account the effects of the emission velocity. Curve (c) is calculated from equation (XI.) as before, but replacing l_g and l_a by $l_g - l_m$ and $l_a - l$ respectively, as is shown in Appendix C, *i. e.*, the position of the potential minimum

is calculated and all distances measured therefrom instead of from the cathode surface.

Fig. 7.

EXPERIMENTAL & CALCULATED CHARACTERISTICS FOR A PLANE TRIODE.



Variation of anode current density i with grid potential V_g ,
at an anode potential of 200 volts.

Curve (a) is calculated from formula (I.).

„ (b) „ „ „ „ „ (XI.).

„ (c) „ „ „ „ „ (XI.) with a correction for initial velocities.

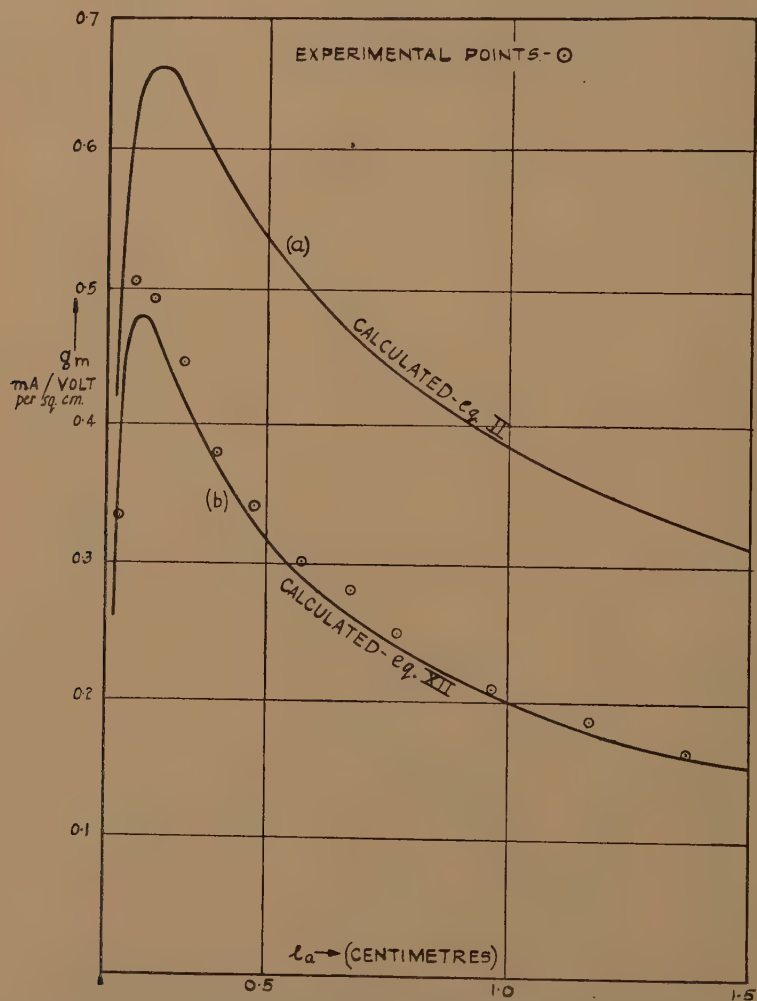
Experimental points are given by ⊙.

$l_g = 0.170$ cm., $l_a = 0.673$ cm., $a = 0.175$ cm., $d = 0.0089$ cm.

The introduction of a filament for the accurate measurement of contact potential would have involved considerable constructional difficulties, and was not attempted.

The effective contact potential between the grid and the cathode was

Fig. 8.



Variation of mutual conductance per unit area g_m with cathode-anode distance l_a at $V_a=200$ volts, $V_g=0$.

$l_g=0.170$ cm., $a=0.175$ cm., $d=0.0089$ cm.

Curve (a) is calculated from equation (II.).

" (b) " " " (XII. a).

Experimental points are given by \odot .

found merely by measuring the applied grid potential for which grid current disappeared (using a microammeter) when a plate current of 10 ma./sq. cm. was flowing, the grid being close to the cathode. This gives the applied grid potential corresponding to the potential minimum rather than to the cathode surface itself. In the particular case of fig. 7 the value of V_m would be about -0.2 volts. All the curves shown are corrected to the true value of cathode potential calculated in this way though the effect of the correction is very small.

The current was always small compared to the saturation current available (about 80 ma./sq. cm.), and in these circumstances l_m can easily be found from the relation given by Langmuir ⁽¹⁰⁾ (see Appendix C).

If the initial velocity correction were made to equation (I.) the disagreement with the experiment would be made appreciably worse instead of better, but when applied to equation (XI.) the agreement with experiment becomes very satisfactory. It is usually more important to be able to calculate the mutual conductance than the current so, having shown the possibility of the latter current calculations will not be further considered in Part I.

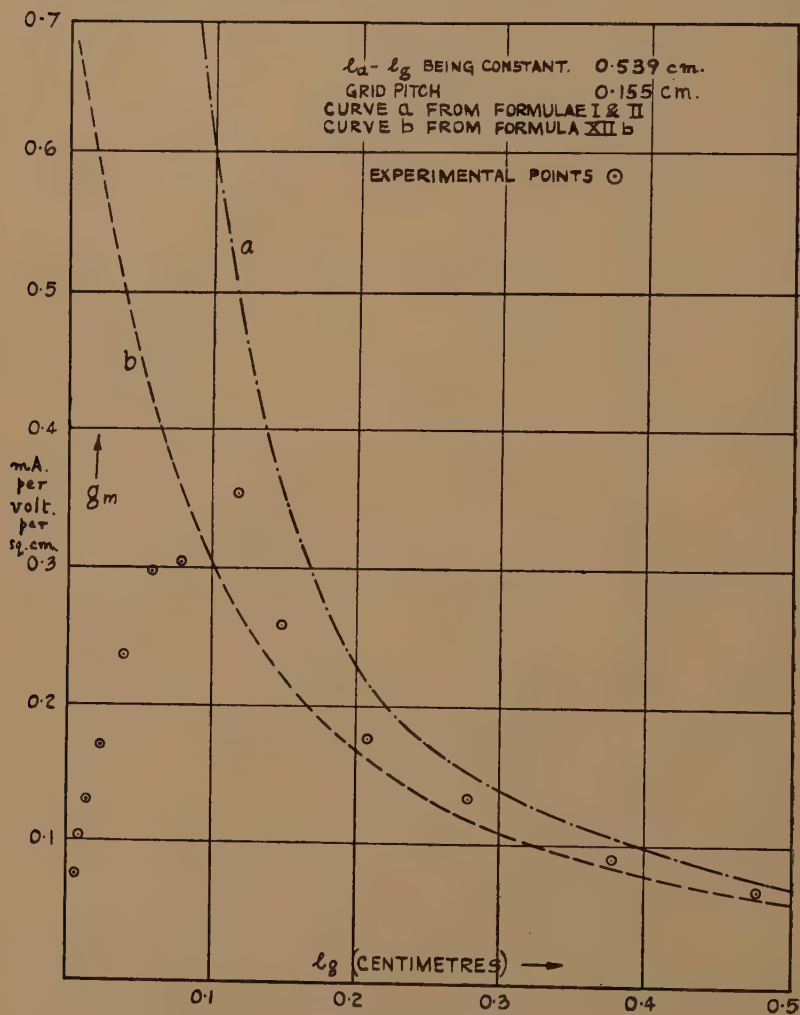
As with amplification factor, the mutual conductance is difficult to measure dynamically, and all measurements were made on anode current-grid voltage characteristics. In fig. 8 is shown the comparison between theory and experiment for the variation of mutual conductance measured at $V_g=0$ with anode distance, the grid distance being kept fixed. In this experiment the correction for initial velocities, always smaller for mutual conductance than for current, has not been made. Making the correction would raise the calculated value of g_m at any given value of anode distance by something of the order of about 2 per cent.; the two curves could not easily be shown in a reproduction. Equation (XII.) derived from the theory proposed above, clearly gives much better agreement than does the more frequently used equation (II.).

Fig. 9 shows the variation of mutual conductance with cathode-grid distance, the grid-anode distance l_a-l_g being kept constant. In this case the mutual conductance is calculated and measured for constant anode current, using equation (XII. *b*). Here it is possible to show the curve, allowing for the emission velocity correction. It can be seen that agreement is good for values of grid-cathode distance greater than the pitch, but that as soon as the grid gets closer to the cathode than this (*i. e.*, by $l_g < a$) the experimentally measured values deviate more and more from the theory. It is particularly striking that the theory shows a continuous increase of g_m as $l_g/a \rightarrow 0$, while the experiment shows a well-marked maximum value at $l_g/a \sim 3/4$.

By a further investigation of the geometrical theory for small values of l_g/a it is possible to account quantitatively for the discrepancy, and to

show why a maximum value of g_m is obtained. This investigation is described in Part II.

Fig. 9.



Variation of mutual conductance per unit area g_m
with cathode-grid distance l_g .

Curves and measurements corresponds to a current of 1 m.a. to F (see fig. 4); that is, $i=1.74$ ma. per sq. cm.

Curve (a) is from equations (I.) and (II.).

" (b) " " equation (XII. b).

Experimental points, \odot .

$l_a - l_g = 0.539$ cm., $a = 0.155$ cm., $d = 0.0050$ cm.

PART II.

Distance between Cathode and Grid small compared to the Grid Pitch.

The formulæ worked out above are all derived upon the assumptions :

1. That the grid wires are thin ; the formulæ hold well as long as the wire diameter is less than 10 per cent. of the grid pitch.
2. That the grid pitch is itself smaller than the distances between the grid and the other two electrodes. When the distance between the grid and the cathode becomes equal to the grid pitch the formulæ will be in error by about 2 per cent. If the grid is far from the cathode the grid-anode distance may be considerably less than the grid pitch without causing serious discrepancies, as it is lack of uniformity of the field at the cathode that causes the major disturbances observable.

It has been mentioned above that formulæ already exist by means of which the penetration factor can be calculated when the grid wires are relatively thick ^{(3), (4)}. It is the purpose of this section to consider the case when the distance between grid and cathode is small. It will be assumed that the grid-anode distance is not small compared to the grid pitch ; the limitation of this will be discussed below.

The analysis is given in Appendix D ; we will give here only the principles on which the calculations were based, together with the results obtained.

In Maxwell's treatise on electricity and magnetism a function which gives the distribution of potential due to a plane charged grid of fine parallel wires is calculated. By adding to this a simple linear function of the distance from the grid plane it is possible to find the function corresponding to a grid between two conducting planes parallel to the grid plane so long as both conducting planes lie at a distance from the grid large compared to the pitch. If, however, the grid lies close to one or other of them the field at the surface of the plane, and hence the charge upon it, is no longer uniform. This means that a linear term can no longer represent the effect of the planes upon the potential distribution. To solve the problem thus set up it is necessary to use the conception of electrostatic images. This shows that the potential distribution due to the grid and to the adjacent plane conductor is exactly the same as would be the appropriate part of the potential distribution between the grid and a similar grid, oppositely charged in the position of the optical image of the original grid in the conducting plane. Alternatively, if two similar but oppositely charged grids lie parallel to each other, with the wires of one exactly opposite to those of the other, there will be an accurately equipotential surface midway between them which can be

identified as one plate of the plane triode to be considered. If a second conducting plane is at a considerable distance the addition to the combined potential functions of the grid and of its imaginary partner of a suitable linear term will adequately represent its contribution to the potential system.

In Appendix D is given the mathematical derivation of the potential function. From this an expression for the electric field at the cathode is calculated. Then it is easy to determine the value of the "electrostatic penetration factor" D_E , which represents the relative dependence of the field at any point of the cathode upon anode and grid voltages.

The general expression obtained for this is

$$D_E = \frac{\left\{ \frac{a}{4\pi l_a} \log_e \left[1 + \frac{\sinh^2 \frac{2\pi l_g}{a}}{\sin^2 \frac{\pi d}{2a}} \right] - \left(\frac{l_g}{l_a} \right)^2 \right\}}{\left\{ \frac{\sinh^2 \left(\frac{2\pi l_g}{a} \right)}{\cosh \left(\frac{2\pi l_g}{a} \right) - \cos \left(\frac{2\pi x}{a} \right)} - \frac{l_g}{l_a} \right\}} - \frac{l_g}{l_a}, \quad (\text{XXVII.})$$

where x is the distance along the cathode measured from a point in the cathode directly below a grid wire.

This is interesting in that it shows no dependence on the values of anode or grid voltage, *i. e.*, the cathode field is still at all points a linear function of V_a and V_g . On the other hand, it is no longer constant over the cathode surface, but shows a periodic variation with x , *i. e.*, as the point considered moves along the cathode.

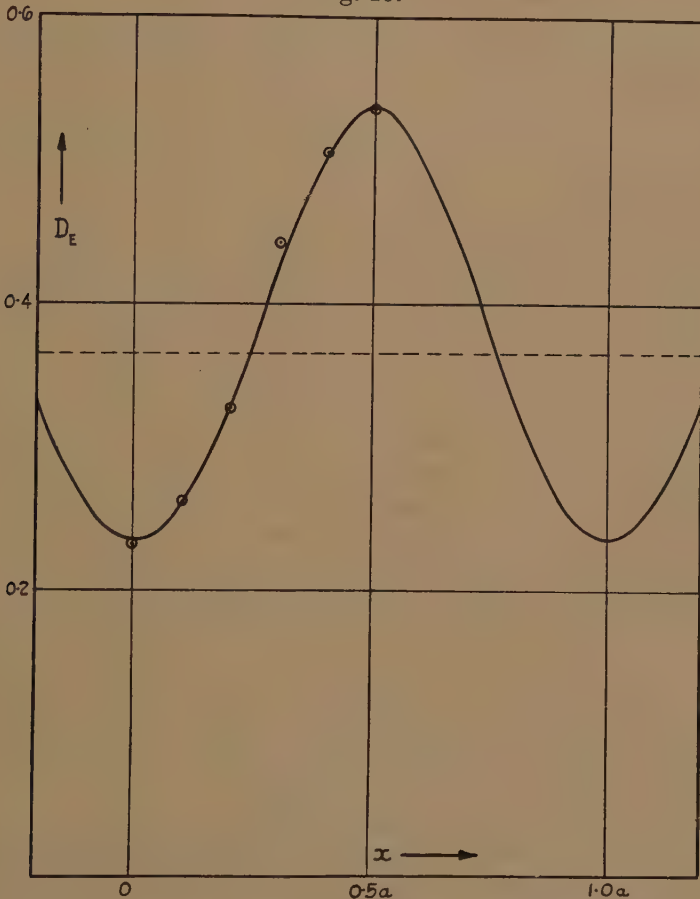
Experiment.

It is not possible to show the variation of penetration factor with x experimentally in a vacuum tube, as in order to do so we should have to know the distribution of current along the cathode. On the other hand, it was felt that it was necessary to get some experimental check of the formula which is of importance in the consideration of variable μ effects in close-spaced valves. It is possible to do this quite accurately using a rubber sheet model such as was originally suggested by Mr. P. B. Moon⁽¹¹⁾, and has been described in some detail by J. R. Pierce and others⁽¹²⁾. A rubber sheet is stretched in a horizontal plane so tightly that it does not sag appreciably under its own weight. Then if points on it are slightly displaced vertically by suitably applied pressure any points on the free parts of the sheet will conform to the equation

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0,$$

where h is the vertical displacement of the point and x, y are rectangular coordinates in a horizontal plane. This equation is of the same form as Laplace's equation for a potential distribution independent of the z axis under space-charge-free conditions, the displacement h taking the place of potential. We can then determine the form of the potential distribution for any system of electrodes whose geometry varies only in two dimensions by applying models of such electrodes to the stretched sheet, their

Fig. 10.



Variation of the electrostatic value of the penetration factor, D_E , at a point on the cathode surface, with the distance x from a point immediately below a grid wire. The curve was calculated from equation (XXVII.) and the experimental points \odot obtained on the rubber sheet model. The dotted line shows the value of D obtained from Schottky's simple formula.

$$a=29.2 \text{ cm.}, \quad d=0.94 \text{ cm.}, \quad l_g=0.4 a, \quad l_a=1.4 a.$$

displacements being proportional to the potentials normally carried by the electrodes. Clearly the slope of the free rubber sheet at any point will then be proportional to the potential gradient at the corresponding point in the electrostatic system to be investigated.

A model of a parallel plane triode was set up in this way. The "field" at the cathode was measured by means of a small piece of mirror lightly attached to the sheet very close to the edge of the cathode. This reflected a beam of light from a fixed source on to a fixed scale. It was then quite easy to measure D_E merely by raising the grid model through a measured height h_g and measuring the distance h_a through which the plate model had to be lowered in order just to bring the light spot back to its original place on the scale. This would correspond in a real valve to increasing the negative bias of the grid by a measured amount and then measuring the increase of anode potential required to maintain the cathode field at its original value.

This was repeated several times, and on plotting h_g against h_a a straight line was obtained of slope D_E . The mirror was used at several points along the cathode surface, and it was thus possible to determine the corresponding variation of D_E (the electrostatic value of penetration factor at a point on the cathode).

In fig. 10 the full line shows the theoretical variation of D_E calculated for a particular case from equation (XXVII.) together with experimental points obtained from the rubber sheet model. The agreement may be regarded as confirmation of equation (XXVII.), though it is possible that mathematicians may rather regard it as proving the ability of the rubber sheet to deal with such problems as this. Certainly the measurements on the sheet give an extremely valuable means of measuring amplification factors in other cases which may not be as easily calculated.

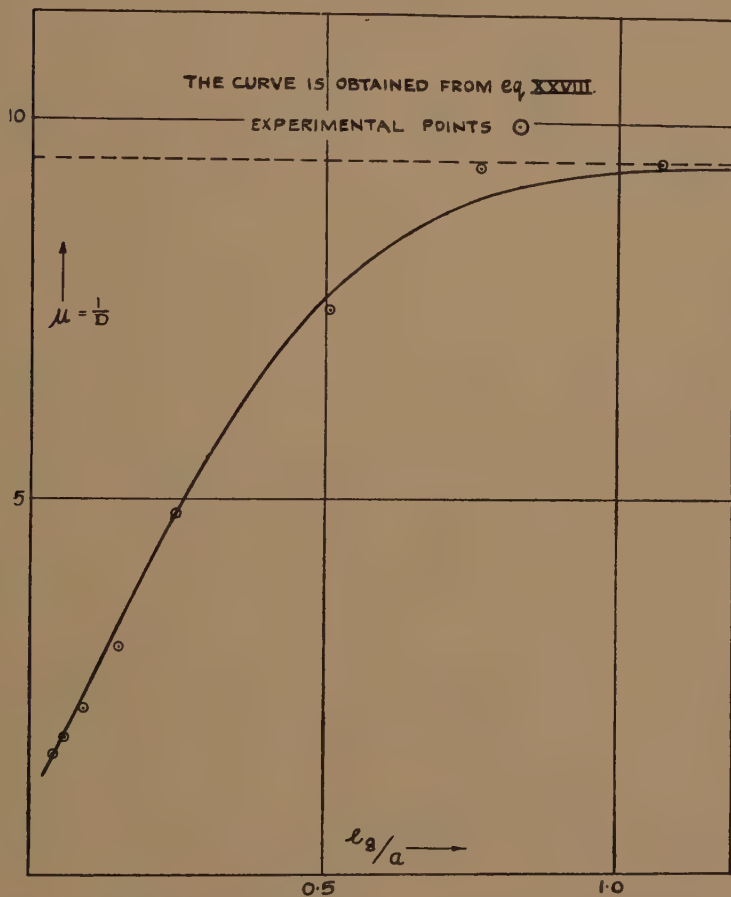
Although it was not possible to measure the variation of D_E at different points of the cathode in the experimental valve, it was possible to measure the particular value D_0 of D_E when $x=(2n+1)a/2$, this being the value which will be obtained for very small currents, as it is clear from general considerations that the current flowing very near to cut-off will be emitted by parts of the cathode midway between grid wires.

Fig. 11 shows the corresponding variation of μ_0 with l_g/a when the distance from grid to anode is kept constant; the full line was calculated from the expression

$$\mu_0 = \frac{1}{D_0} = \frac{4\pi \left[\frac{l_g}{a} - \frac{l_a}{a} \tanh\left(\frac{\pi l_g}{a}\right) \right]}{\frac{4\pi}{a} l_g \tanh\left(\frac{\pi l_g}{a}\right) - \log_e \left[1 + \frac{\sinh^2\left(\frac{2\pi l_g}{a}\right)}{\sin^2 \frac{\pi d}{2a}} \right]}, \quad (\text{XXVIII.})$$

to which expression (XXVII.) reduces when $x=a/2$, and the experimental points were obtained by plotting μ against current and extrapolating to zero.

Fig. 11.



Variation of the penetration factor D and amplification factor μ at small values of the grid cathode distance l_g . The curve is calculated from equation (XXVII.) and the points \odot are measured experimentally for small currents.

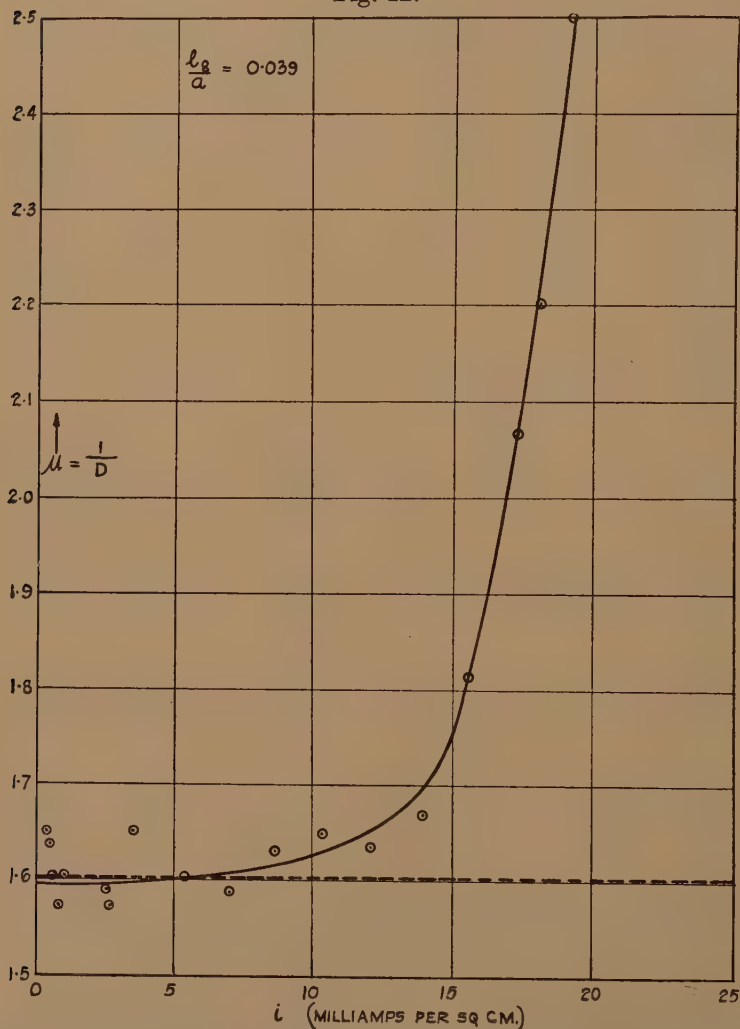
$$(l_a - l_g^*) = 0.539 \text{ cm.}, \quad a = 0.155 \text{ cm.}, \quad d = 0.0050 \text{ cm.}$$

It seems then that the expression gives accurate values down to very small values of l_g/a . For practical reasons it has not yet been tried for

very small grid-anode distances ; using the rubber sheet model agreement was still obtained for

$$\frac{l_a - l_g}{a} = 0.4 \quad \text{and} \quad \frac{l_g}{a} = 0.4.$$

Fig. 12.

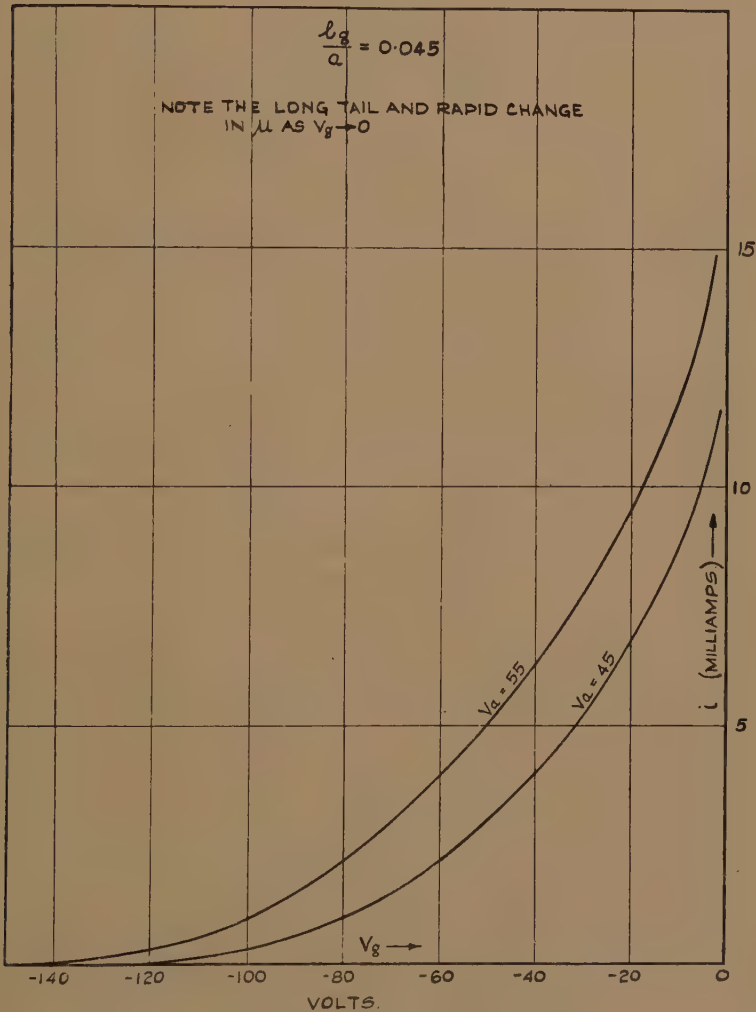


Variation of the penetration factor and amplification factor with anode current density when the grid is very close to the cathode. The dotted line represents the value obtained for the cut-off value from equation (XXVIII); the points and curve are experimental. Schottky's formula gives $\mu = 1/D = 9.4$. Anode potential 140-180 volts.

$l_g = 0.006$ cm., $l_a = 0.545$ cm., $a = 0.155$ cm., $d = 0.0050$ cm.

In order to calculate anode current and mutual conductance we need to know the variation of penetration factor with current. It can be seen from fig. 10 that D (the average value as measured) will not begin to change very much until emission has begun over an appreciable part of the cathode surface; furthermore it will be the mean value of D_E

Fig. 13.



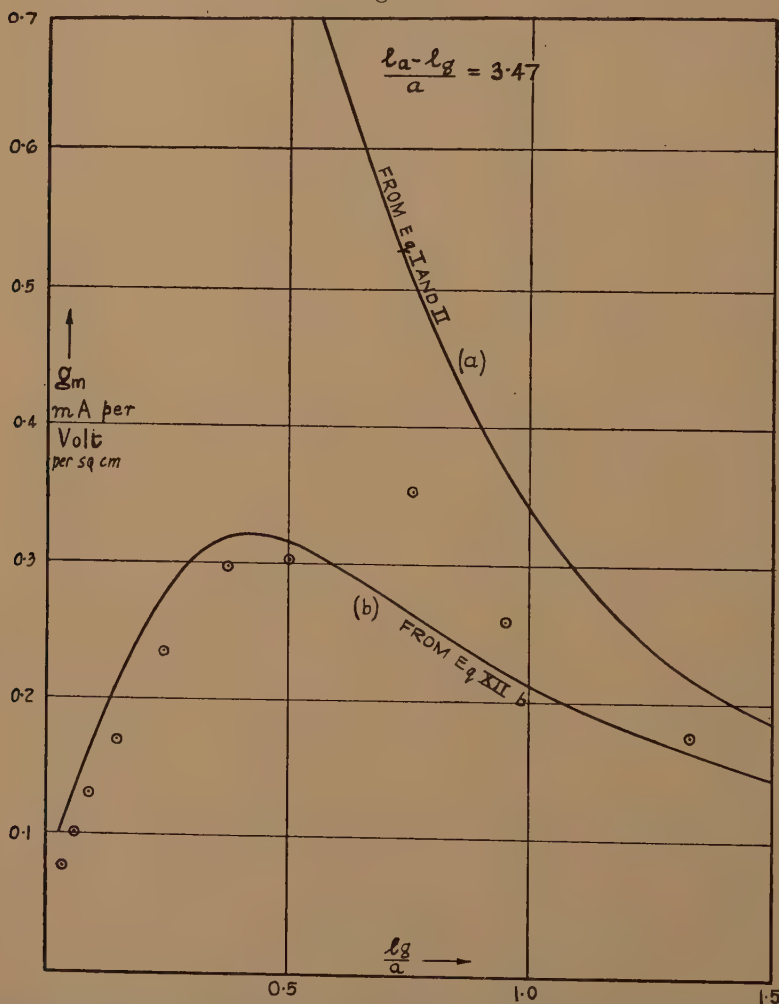
Experimental anode current-grid voltage characteristics under conditions of severe inselbildung. Note the long tail and the rapid change of the penetration factor as $V_g \rightarrow 0$.

$$l_g = 0.007 \text{ cm.}, \quad l_a = 0.162 \text{ cm.}, \quad a = 0.155 \text{ cm.}, \quad d = 0.0050 \text{ cm.}$$

along the emitting part of the surface weighted according to the current density; this being greatest when D_E is greatest will help to prevent a rapid change of D with small currents.

Fig. 12 shows how the experimentally measured amplification factor changed with current with the grid very close to the cathode when l_g/a was 0.039. The dotted line shows the theoretical value, $\mu_0 = 1/D_0$, for

Fig. 14.



Variation of mutual conductance g_m per unit area with cathode grid distance measured as a fraction of the grid pitch, l_g/a . Theoretical curves, experimental points at a current density $i = 1.74$ ma. per sq. cm.

$$(l_a - l_g) = 0.539 \text{ cm.}, \quad a = 0.155 \text{ cm.}, \quad d = 0.0050 \text{ cm.}$$

zero current. Up to a current density of 10 ma./sq. cm. practically no change takes place from the cut-off value μ_0 . Beyond this μ begins to go up very rapidly as $V_g \rightarrow 0$, when current begins to flow from parts of the cathode immediately below grid wires, as would be expected. The sharpness of the "elbow" in fig. 12 depends on the values of V_a and V_g . The type of characteristic obtained at such close spacing is shown in fig. 13. For a given value of anode current a high anode potential and large negative grid bias will give a much sharper elbow than will a low anode potential and correspondingly lower grid bias. The reason for this is clear at once from consideration of the way in which the current will be distributed along the cathode surface in the two cases.

For fairly small currents then it should be possible to get a close approximation to the measured mutual conductance from equation (XII.), using the electrostatic value for the penetration factor in the calculation, taken from equation (XXVIII.). Fig. 14 shows a curve calculated in this manner for the variation of mutual conductance with l_g , assuming that 1 ma. flows in the centre section of the anode used in the experiment. The calculation is carried out for the case shown in fig. 9. The curve shows the well-marked maximum found experimentally; this is clearly demanded by the fact that when l_g/a is small g_m is proportional to $D^{-3/2}$ and D itself is roughly inversely proportional to l_g/a *. The values of l_g/a , for which the mutual conductance is a maximum, for constant anode-grid distance and in some other given conditions have been calculated. The calculations are not given here, as the expressions obtained are somewhat unwieldy and are of limited importance. The maxima are flat, and the exact positions found in practice depend on the emission velocities of the electrons. The experimental points shown in fig. 9 are again put in, and it is clear that by taking account of the variation of the penetration factor with grid-cathode distance a close approximation to the measured mutual conductance can be obtained. If initial velocities are allowed for a closer approximation still can be obtained. This is clearly exemplified by fig. 15, which was taken for a smaller value of grid-anode distance in which the effect of emission velocity was greater and, since the saturation current was higher, more easily estimable.

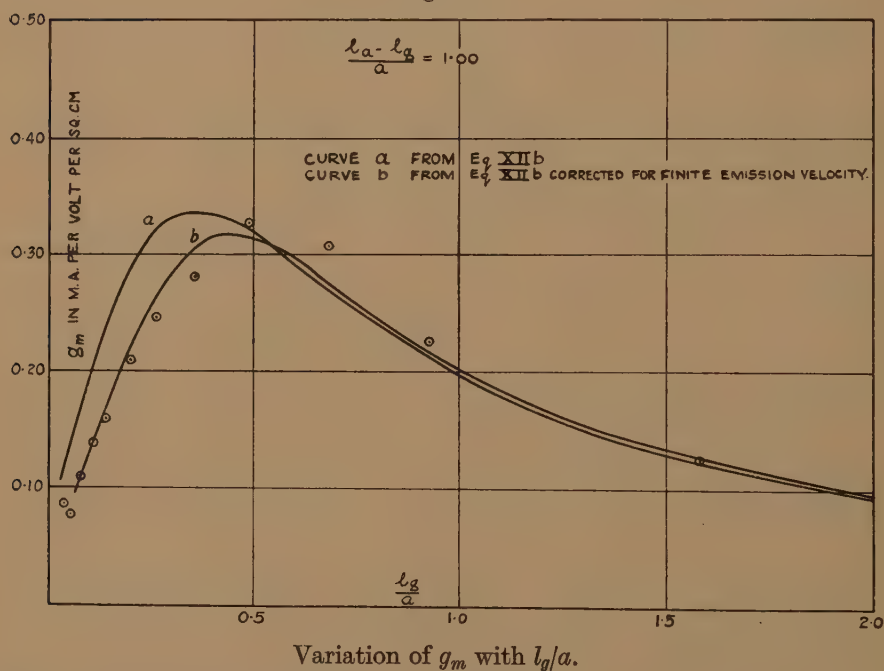
We may conclude then that if allowance is made for the increase of the penetration factor at close spacings, the formulæ (XI.), (XII.) which have been developed make it possible to calculate anode current or mutual conductance to a very fair accuracy for any valve in which the electrodes may be regarded as plane, so long as the current density is not too high. Accurate calculations of the upper parts of the characteristics for close

* This can be seen from equations (XII. b) and (XXVII.); the way in which the penetration factor varies with cathode-grid distance can be seen from fig. 11.

spacings have not yet been made ; work is still being done on this problem. It should be noted that for valves with large amplification factors the calculated value of anode current and mutual conductance will be approximate, owing to the big difference made to the effective voltage by contact potentials etc., unless this can be measured by observation of the grid potential at which grid current disappears. This is impossible when it is desired to carry out the calculations in advance. It is often useful, then, to decide in advance the anode current to be used at the working point and to use the formula (XII. *b*)

$$g_m = \frac{26.4 \sqrt[3]{i} \times 10^{-3} \text{ amps. per volt per sq. cm. when } i \text{ is in}}{l_g^{4/3} + D l_a^{4/3}} \text{ amps. per sq. cm.,}$$

Fig. 15.



Curve (a) from equations (XII. *b*) and (XXVIII.).

„ (b) „ „ (XII. *b*) „ (XXVIII.) corrected for finite emission velocity.

Experimental points, \odot .

$(l_a - l_g) = 0.155$ cm., $a = 0.155$ cm., $d = 0.0050$ cm., $i = 1.74$ ma. per sq. cm.

derived from (XI.) and (XII. *a*) by elimination of V_a and V_g to determine the mutual conductance. The exact grid bias necessary is then found after the valve has been made up.

The work here described has been carried out in the Valve Laboratory of Standard Telephones and Cables, Limited, Woolwich, to whom I am indebted for permission to publish these results. I wish to acknowledge also my grateful thanks to Mr. W. T. Gibson, the Chief Valve Engineer, for the long hours which he spent in giving invaluable suggestions as to the best way of representing the results obtained; to Dr. D. H. Black, Chief of the Valve Laboratory, for his continuous advice and encouragement; to R. N. Hall and D. P. R. Petrie for their help with the work itself; and to W. R. Hindle for his skilful and accurate work in the construction of the experimental apparatus.

Finally, I would like to thank Mr. J. A. Ratcliffe for a very helpful discussion, as a result of which the theory of Part I. was considerably clarified.

APPENDIX A.

<i>Symbol.</i>	<i>Meaning.</i>
$a.$	Grid pitch.
A, α , B.	Used as constants.
$\beta^2.$	Langmuir's function for a cylindrical diode. See ref. (7).
$d.$	Diameter of a grid wire.
$D = \frac{\partial i}{\partial V_a} / \frac{\partial i}{\partial V_g}.$	Penetration factor (reciprocal of amplification factor).
$D_E = \frac{\partial}{\partial V_a} \left(\frac{\partial V}{\partial y} \right)_c \bigg/ \frac{\partial}{\partial V_g} \left(\frac{\partial V}{\partial y} \right)_c.$	Penetration factor, electrostatically calculated.
$D_0.$	The value of D_E midway between wires.
$e.$	Electronic charge.
$\epsilon.$	Base of natural logarithms.
$\eta.$	The function $\frac{e}{kT} (V - V_m).$
$\eta_c.$	The value of η at the cathode; its value is $\log_e \left(\frac{i_{sat}}{i} \right).$ See ref. (10).
$g_m.$	Mutual conductance per unit area in formulæ for plane electrodes; per unit length for cylindrical electrodes.
i or $i_l.$	Space current per unit area (plane) or unit length (cylindrical).
$i_{sat}.$	Space current density in saturation conditions.
$I_F, I_E.$	Experimentally measured currents to the centre anode and outer guard ring respectively. See fig. 4.

k .	Boltzmann's Constant.
ξ .	The function
	$4(l-l_m)\left(\frac{\pi}{2kT}\right)^{3/4}m^{1/4}(ei)^{1/2}=2L(l-l_m).$
ξ_c .	The value of ξ at the cathode. See ref. ⁽¹⁰⁾ .
l .	Distance of a point from the cathode.
l_a, l_g, l_m .	Distance of anode, grid, and potential minimum respectively from the cathode.
l_D .	Distance between the electrodes of the equivalent diode.
L .	Defined by $L=2\left(\frac{\pi}{kT}\right)^{3/4}m^{1/4}\sqrt{ei}$.
λ .	Charge per unit length of grid wires.
m .	Electronic mass.
$\mu=\frac{\partial i}{\partial V_g}\bigg \frac{\partial i}{\partial V_a}$.	Amplification factor.
μ_0 .	Reciprocal of D_0 . (Amplification factor at cut-off.)
n .	An integer.
r_a, r_g, r_c .	Radii of anode, grid, and cathode respectively of a cylindrical triode.
T .	Absolute temperature.
V .	Potential at any point.
V_a, V_g, V_m .	Potentials of anode, grid, and potential minimum with respect to the cathode.
x, y .	Rectangular coordinates with origin in the cathode, the axis of x being in the plane of the cathode perpendicular to the wires and the axis of y passing through the centre of a wire and increasing towards the anode.
$\left(\frac{\partial V}{\partial y}\right)_c$.	The potential gradient at the cathode surface.

APPENDIX B.

References.

- (1) J. M. Miller, P. I. R. E. viii. p. 64 (1920).
- (2) R. W. King, Phys. Rev. xv. p. 256 (1920).
- (3) F. B. Vogdes and F. R. Elder, Phys. Rev. xxiv. p. 683 (1924).
- (4) F. Ollendorf, 'Elektrotech und Maschinebau,' lii. p. 585 (1934).
- (5) See, for example, Benjamin, Cosgrove, and Warren, I. E. E. (April 1937).
- (6) Child, Phys. Rev. xxxii. p. 142 (1911).

- (7) Langmuir, Phys. Rev. ii. p. 450 (1913).
- (8) Chaffee, 'The Theory of Thermionic Vacuum Tubes,' pp. 146-7, 1st edition.
- (9) Dow, 'Fundamentals of Engineering Electronics,' equations (82) and (111).
- (10) Langmuir, Phys. Rev. xxi. p. 419 (1923).
- (11) P. B. Moon and M. L. Oliphant, Proc. Camb. Phil. Soc. xxv. p. 461 (1929).
- (12) J. R. Pierce, Bell Labs. Record, xvi. p. 305 (1938); P. H. J. A. Kleyne, Philips Techn. Rev. ii. p. 338 (1937).
- (13) W. E. Benham, P. I. R. E. xxvi. p. 1093 (1938).

APPENDIX C.

As a first approximation we may correct for the effect of initial velocity of emission from the cathode by measuring all distances and voltages from the potential minimum outside the cathode rather than from the cathode itself. Langmuir ⁽¹⁰⁾ has shown that this potential minimum always exists if the current is unsaturated, and has shown how in a plane diode its distance from the cathode l_m and depth V_m can be calculated for any given current density if the saturation current density i_{sat} is known. He assumes the emitted electrons to have a Maxwellian distribution of velocities corresponding to the temperature of the cathode. For convenience he introduces the non-dimensional variables

$$\left. \begin{aligned} \eta &= \frac{e}{kT} (V - V_m), \\ \xi &= 4 \left(\frac{\pi}{2kT} \right)^{3/4} m^{1/4} \cdot \sqrt{ei} (l - l_m) = 2L(l - l_m). \end{aligned} \right\} \dots (XV.)$$

The relation between η and ξ is then independent of the properties of any particular diode, and tables are given in the paper previously referred to from which either can be found if the other is known. Using suffices to indicate the points considered we have $\eta_m = 0$, and Langmuir shows that

$$\eta_c = \log_{\epsilon} \frac{i_{sat}}{i}.$$

Then

$$V_m = - \frac{kT}{e} \log_{\epsilon} \frac{i_{sat}}{i},$$

or in practical units

$$V_m = - \frac{T}{11,600} \log_{\epsilon} \frac{i_{sat}}{i} \dots \dots \dots (XVI.)$$

To calculate the value of l_m we use the tables mentioned above to find $-\xi'_c$, since η_c can be found when i_{sat} is known.

Then

$$-\xi_c = 2Ll_m,$$

and in practical units

$$l_m = \frac{-\xi_c}{2.90 \times 10^4 T^{-3/4} \sqrt{i}} = \frac{\alpha}{\sqrt{i}} \text{ say. } \dots \dots \dots (XVII.)$$

If voltages and distances are measured from the potential minimum then, from (XI.), in the plane case,

$$i = \frac{2.34 \times 10^{-6} (V_g + DV_a - V_m)^{3/2}}{[(l_g - l_m)^{4/3} + D(l_a - l_m)^{4/3}]^{3/2}} \quad \text{amps per unit area. (XVIII.)}$$

This expression, in spite of its approximate nature, would give a very complex formula for the mutual conductance as it stands, since l_m and V_m both vary with i . If i is not a large fraction of i_{sat} , however, we can make some further assumptions before differentiating. For large values of η the rate of variation of $-\xi$ with η is small. If this is neglected we have, as in (XVII.), that l_m is inversely proportional to \sqrt{i} , α being assumed constant for the purpose of differentiation. When i/i_{sat} becomes fairly large, α is no longer even approximately constant, but, on the other hand, the whole correction required is very small. The effect of the variation of V_m is very small except at very small currents indeed.

Then we have, from equation (XVIII.),

$$\begin{aligned} V_g + DV_a - V_m &= 5690 \cdot i^{2/3} \left[\left(l_g - \frac{\alpha}{\sqrt{i}} \right)^{4/3} + D \left(l_a - \frac{\alpha}{\sqrt{i}} \right)^{4/3} \right]; \\ \therefore \frac{\partial V_g}{\partial i} &= \frac{2}{3} \cdot 5690 \cdot i^{-1/3} \left[\left(l_g - \frac{\alpha}{\sqrt{i}} \right)^{4/3} + D \left(l_a - \frac{\alpha}{\sqrt{i}} \right)^{4/3} \right] \\ &\quad + 5690 \cdot i^{2/3} \left[\frac{2\alpha}{3} \left(l_g - \frac{\alpha}{\sqrt{i}} \right)^{1/3} i^{-3/2} + \frac{2}{3} D\alpha \left(l_a - \frac{\alpha}{\sqrt{i}} \right)^{1/3} i^{-3/2} \right], \end{aligned}$$

which gives us finally

$$g_m = g_{m1} \left[1 + l_m \frac{(l_g - l_m)^{1/3} + D(l_a - l_m)^{1/3}}{(l_g - l_m)^{4/3} + D(l_a - l_m)^{4/3}} \right]^{-1}, \quad \dots \quad \text{(XIX.)}$$

where g_{m1} is the value of g_m obtained from equation (XII. *b*) by replacing l_g , l_a by $(l_g - l_m)$, $(l_a - l_m)$ respectively for a given current. The factor in the bracket is usually very close to 1.

In the case of a valve in which the electrodes are concentric cylinders the correction for initial velocities is considerably smaller ⁽¹⁰⁾, and may therefore be neglected.

APPENDIX D.

If we have a grid at a distance l_g from a conducting plane cathode we can consider the potential distribution as being that due to a pair of oppositely charged grids at $y = \pm l_g$. ($y = 0$ at the cathode). The expression

$$V = -\lambda \log_e 2 \left[\cosh \left(\frac{2\pi y}{a} \right) - \cos \left(\frac{2\pi x}{a} \right) \right]$$

represents the potential at any point (x, y) due to a plane grid of fine equidistant parallel wires with charge per unit length λ , in the plane $y = 0$, the axis of x being taken perpendicular to the grid wires and the origin being at the centre of a wire.

Then for two grids at $y = \pm l_g$ we should have

$$V = -\lambda \log_e \left[\frac{\cosh \frac{2\pi y - l_g}{a} - \cos \frac{2\pi x}{a}}{\cosh \frac{2\pi y + l_g}{a} - \cos \frac{2\pi x}{a}} \right] \quad \dots \quad (\text{XX.})$$

If we have an anode parallel to these at a considerable distance at $y = l_a$ we shall obtain the correct potential at any point by adding $By + C$ to the potential V , where B and C are constants. At the cathode, where $y = 0$, the potential then reduces to

$$V_1 = C. \quad \dots \quad (\text{XXI.})$$

At the grid $y = l_g$, $x = d/2$ (the wire radius), and we have

$$V_2 = -\lambda \log_e \left[\frac{1 - \cos \left(\frac{\pi d}{a} \right)}{\cosh \left(\frac{4\pi l_g}{a} \right) - \cos \left(\frac{\pi d}{a} \right)} \right] + Bl + C;$$

$$\therefore V_2 = \lambda \log_e \left[1 + \frac{\sinh^2 \left(\frac{2\pi l_g}{a} \right)}{\sin^2 \left(\frac{\pi d}{2a} \right)} \right] + Bl_g + C, \quad \dots \quad (\text{XXII.})$$

and at the anode

$$V_3 = \lambda \log_e \left[\frac{\cosh \left(\frac{2\pi}{a} \cdot \overline{l_a + l_g} \right)}{\cosh \left(\frac{2\pi}{a} \cdot \overline{l_a - l_g} \right)} \right] + Bl_a + C$$

$$= \lambda \cdot \frac{4\pi}{a} l_g + Bl_a + C \quad \dots \quad (\text{XXIII.})$$

if l_a/a is fairly large.

From the last four equations then we have

$$\left. \begin{aligned} V_g &= \lambda \log_e \left[1 + \frac{\sinh^2 \left(\frac{2\pi l_g}{a} \right)}{\sin^2 \left(\frac{\pi d}{2a} \right)} \right] + Bl, & (a) \\ V_a &= \frac{4\pi l_g}{a} \cdot \lambda + Bl_a, & (b) \\ V &= \lambda \log_e \left[\frac{\cosh \left(\frac{2\pi}{a} \cdot \overline{y + l_g} \right) - \cos \left(\frac{2\pi x}{a} \right)}{\cosh \left(\frac{2\pi}{a} \cdot \overline{y - l_g} \right) - \cos \left(\frac{2\pi x}{a} \right)} \right] + By, & (c) \end{aligned} \right\} \quad (\text{XXIV.})$$

where V is now measured from the cathode.

From these we can find the potential at any point by elimination of B and λ in terms of V_a and V_g . Here, however, the field at the cathode,

$\left(\frac{\partial V}{\partial y}\right)_c$, is required. This is given, from (XXIV. c), by

$$\left(\frac{\partial V}{\partial y}\right)_c = \frac{\frac{4\pi}{a} \lambda \sinh\left(\frac{2\pi}{a} l_g\right)}{\cosh\left(\frac{2\pi}{a} l_g\right) - \cos\left(\frac{2\pi x}{a}\right)} + B. \quad \text{(XXV.)}$$

Eliminating λ and B from (XXV.), (XXIV. a), and (XXIV. b),

$$\begin{aligned} \left(\frac{\partial V}{\partial y}\right)_c = \frac{V_a}{l_a} + \frac{(l_g V_a - l_a V_g) \frac{4\pi}{a}}{\frac{4\pi}{a} l_g^2 - l_a \log_e \left[1 + \frac{\sinh^2\left(\frac{2\pi}{a} l_g\right)}{\sin^2\left(\frac{\pi d}{2a}\right)} \right]} \\ \times \left[\frac{l_g}{l_a} + \frac{\sinh\left(\frac{2\pi}{a} l_g\right)}{\cosh\left(\frac{2\pi}{a} l_g\right) - \cos\left(\frac{2\pi x}{a}\right)} \right]. \quad \text{(XXVI.)} \end{aligned}$$

This will vary with x until

$$\cosh\left(\frac{2\pi}{a} l_g\right) \gg 1.$$

When $l_g/a = 1$ the variation will not be very great, as $\cosh 2\pi$ is about 275, but for smaller values the variation may become considerable unless $(l_g V_a - l_a V_g)$ is small.

The relative effectiveness of the grid and anode voltages then varies along the cathode. We shall define as "electrostatic penetration factor," D_E , at any point the ratio

$$\frac{\partial}{\partial V_a} \left(\frac{\partial V}{\partial y} \right)_c / \frac{\partial}{\partial V_g} \left(\frac{\partial V}{\partial y} \right)_c.$$

The value of D_E can then be found from (XXVI.) to be given by

$$D_E = \frac{\left\{ \frac{a}{4\pi l_a} \log_e \left[1 + \frac{\sinh^2\left(\frac{2\pi}{a} l_g\right)}{\sin^2\left(\frac{\pi d}{2a}\right)} \right] - \frac{l_g^2}{l_a^2} \right\}}{\left\{ \frac{\sinh\left(\frac{2\pi}{a} l_g\right)}{\cosh\left(\frac{2\pi}{a} l_g\right) - \cos\left(\frac{2\pi x}{a}\right)} - \frac{l_g}{l_a} \right\}} - \frac{l_g}{l_a}, \quad \text{(XXVII.)}$$

which again clearly varies along the cathode surface, giving a maximum value between two wires. In fig. 10 is shown the form of variation for the particular case

$$\frac{l_g}{a} = 0.4, \quad \frac{l_a}{a} = 1.4.$$

The maximum value D_0 is given by

$$D_0 = \frac{1}{4\pi} \frac{\frac{4\pi}{a} l_g \tanh\left(\frac{\pi l_g}{a}\right) - \log_e \left[1 + \frac{\sinh^2\left(\frac{2\pi}{a} l_g\right)}{\sin^2\left(\frac{\pi d}{2a}\right)} \right]}{\frac{l_g}{a} - \frac{l_a}{a} \tanh\left(\frac{\pi l_g}{a}\right)}, \quad (\text{XXVIII.})$$

which for small values of l_g/a reduces to

$$D_0 = \frac{\log_e \left(1 + \frac{l_g^2}{a^2} \right)}{\frac{2\pi}{a} l_g \left(\frac{\pi l_a}{a} - 1 \right)}. \quad \dots \dots \dots (\text{XXIX.})$$

The theoretical and experimental variation of D_0 with l_g/a for values of the latter less than unity is shown in fig. 11 for a constant $\frac{l_a - l_g}{a} = 1$ and $\frac{d}{a} = 0.032$. As d/a increases D_0 falls away more rapidly when l_g/a decreases, and hence when a grid close to the cathode is to be used it is advantageous to use the finest possible wire to avoid as far as possible the effects of variation with current of D .

The Valve Laboratory,
Standard Telephones and Cables, Ltd.,
North Woolwich, E. 16.

LXIX. *The Evidence for a Superlattice in the Nickel-iron Alloy Ni₃Fe* *.

By P. LEECH, B.Sc., and C. SYKES, D.Sc.

[Received April 21, 1939.]

I. *Synopsis.*

THE nickel-iron alloys containing from 40–90 per cent. nickel by weight possess very interesting magnetic properties, and in particular the alloy containing 78·5 per cent. nickel by weight, sometimes referred to as “permalloy,” has a very high permeability $\sim 100,000$ in small magnetic fields. The magnetic properties of this alloy can be varied to a marked extent by heat-treatment, and many attempts have been made to correlate the changes in magnetic properties with the changes in structure, internal strain, etc., which are normally associated with variations in heat-treatment.

As the composition of permalloy is very near to the composition Ni₃Fe (76·0 per cent. by weight of nickel), it has been suggested that the alloy might undergo an order-disorder transformation on cooling. In this event the degree of order existing in the alloy at room temperature may vary with the rate of cooling and give rise to the observed variations in magnetic properties. On the other hand, the behaviour of the alloy has been interpreted in terms of the theory that for some critical cooling rate the internal thermal strain left in the alloy will just balance that set up due to magneto-striction on magnetization, and in this way enable high permeabilities to be obtained.

A number of investigations have been carried out to establish the existence of an order-disorder transformation in Ni₃Fe. Dahl ⁽¹⁾ has investigated the effect of cold work on the electrical resistivity of Ni₃Fe and has compared the results with the effect of similar operations on the alloy Cu₃Au. In the case of Cu₃Au as quenched, *i. e.*, when the alloy is in the disordered state, the resistivity of the alloy is only increased by about 2 per cent by severe cold working. If, however, the alloy is cold worked from the ordered state then the resistivity is increased very markedly by cold working, *i. e.*, upwards of 50 per cent., and finally attains the same value as in the cold worked disordered alloy. The

* Communicated by Prof. W. L. Bragg, F.R.S.

various stages in the breakdown of the ordered state on cold working can be followed by X-ray methods:—the superlattice lines become weaker and finally disappear, showing the alloy to be disordered. For the Ni₃Fe alloy Dahl found that the increase in resistivity with increased cold working was greater than would be expected if the alloy was a solid solution in which the iron and nickel atoms were distributed at random. The maximum effect was found after the alloy had been annealed for 36 hours at 420° C., when the resistivity increases by some 35 per cent. after a 70 per cent. reduction in cross-sectional area. In the quenched state the resistivity changes by about 15 per cent. after similar cold working. Dahl explains the increase in resistivity of the quenched alloys as being due to a partial ordering which might occur during the actual quenching process. The similarity in behaviour of the alloys Cu₃Au and Ni₃Fe indicated that Ni₃Fe may exist in either the ordered or disordered state but gave no conclusive proof.

When an alloy can exist in either an ordered or disordered state, then the lattice spacing as measured by X-ray methods is found to be different in the two cases. Bradley, Jay, and Taylor ⁽²⁾ have measured the lattice spacing of the iron-nickel alloys. The specimens were examined in four conditions: slowly cooled, and quenched from 700° C., 800° C., and 900° C. In the region of composition around Ni₃Fe it was found that the lattice spacings of the quenched alloys varied in a somewhat erratic fashion and were not readily reproducible. They were, in general, higher than those for the corresponding alloys in the slowly cooled condition. No superlattice lines were visible on the X-ray photographs. In another paper on the ternary system Fe, Ni, Al, Bradley and Taylor ⁽³⁾ have reported that if a small amount of the nickel in Ni₃Fe is replaced by aluminium, as, for example, in Ni₁₄Fe₅Al, then superlattice lines appear on X-ray photographs of the slowly cooled alloy taken with cobalt K α radiation. They concluded that it might be possible to form a superlattice in Ni₃Fe provided the alloy was annealed for a long period below 600° C. Since the experimental work described in the present paper was commenced further papers have appeared. Kaya ⁽⁴⁾ has examined the specific heat temperature curves of Ni₃Fe in the as quenched condition and after slow cooling from 700° C. The resulting curves gave strong evidence for the conclusion that the slowly cooled alloy was ordered and the quenched alloy disordered. Using the K α radiation of nickel he attempted to establish the existence of a superlattice by X-ray methods, but was unsuccessful.

Haworth ⁽⁵⁾ has carried out a very careful X-ray investigation using the K β radiation of iron, but was unable to find any evidence of a superlattice, although control experiments using β brass indicated that the technique used was entirely adequate.

It appears, therefore, that the experimental evidence for the existence of a superlattice in slowly cooled Ni_3Fe is conflicting. The present paper describes further experimental evidence dealing with this problem: this confirms the specific heat data of Kaya and provides for the first time definite X-ray evidence of a superlattice in Ni_3Fe .

II. Experimental.

The alloy used in the investigations was prepared from a pure Swedish iron, Fe 99.96 per cent., and Mond Nickel of high purity, Fe+Ni 99.95 per cent. The alloying was carried out in a high frequency furnace at a pressure of 10^{-3} mm. The ingot was forged to 1 in. diameter (a reduction of about 50 per cent. in area) and annealed *in vacuo* at 900°C . The specific heat specimen, magnetic test pieces, and powder for X-ray examination were all prepared from the same ingot. The chemical composition of the alloy was Fe 24.91 per cent., Ni 75.04 per cent., Al less than .02 per cent., *i. e.*, 74.3 per cent. Ni by atoms.

III. Specific Heat Measurements.

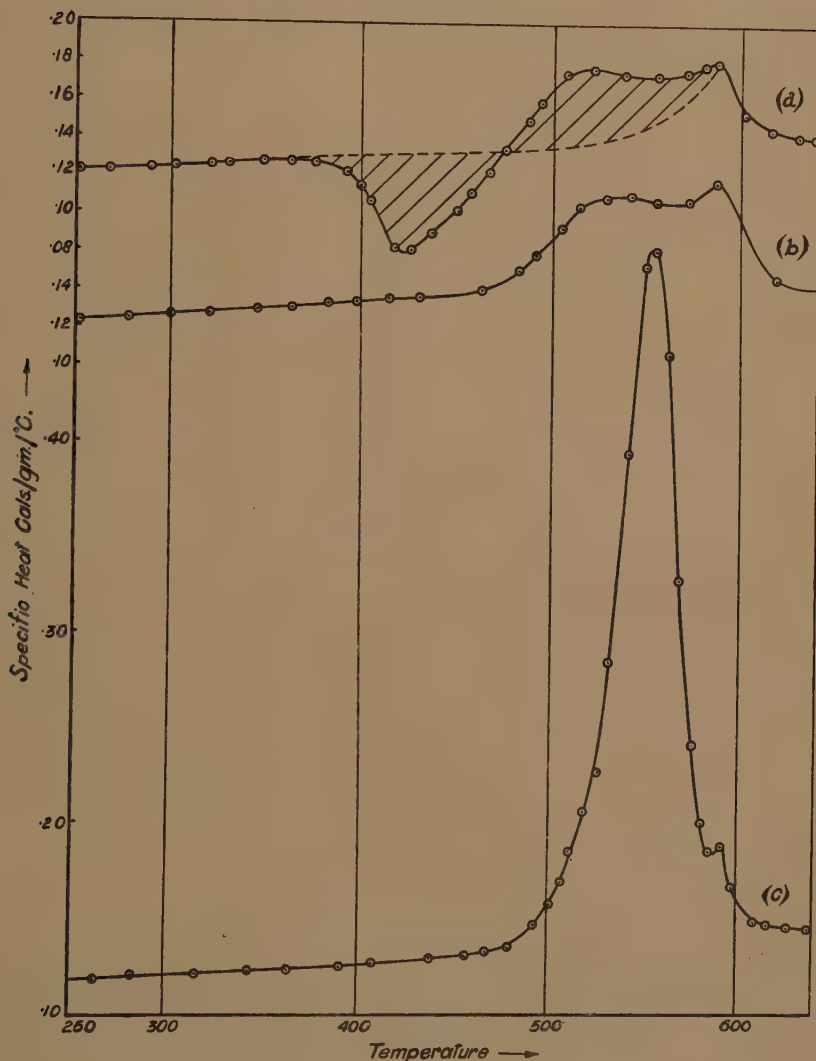
The technique used was identical with that employed in the investigation of other order-disorder transformations ⁽⁶⁾, and has been described in full elsewhere. Reference will be made, therefore, only to the experimental results. The specific heat measurements were made continuously whilst the temperature of the specimen was increased at an approximately constant rate of 2°C . per minute.

Fig. 1 gives a number of specific heat temperature curves which were obtained on the specimen after it had been subjected to different heat treatments. Fig. 1 (a) was obtained on the alloy after quenching from 700°C . in water. The specific heat rises steadily to about 380°C ., then drops to a minimum at 425°C . At higher temperatures two maxima are observed, one at 525°C . and the other at 590°C . This curve is more complicated than those of ordinary ferromagnetic materials such as nickel and iron; in such cases the specific heat rises steadily until the temperature approaches the Curie point, then there is a sharp maximum and a rapid fall to a fairly steady value, *i. e.*, there is only one maximum and no minimum.

From magnetic measurements described in a subsequent section the maximum at 590°C . is to be associated with the Curie point of the alloy; the additional complications must be ascribed to transformations other than the normal magnetic transformation. The experimental evidence indicates that the second transformation is almost certainly an order-disorder transformation; the minimum in fig. 1 (a) is produced by the partial ordering of the quenched disordered alloy, whilst the first maximum at 525°C . is the temperature at which the partially ordered alloy

returns to the disordered state. Fig. 1 (b) refers to the alloy as cooled at $1^\circ\text{C. per minute}$ from 650°C. The minimum has disappeared, presumably

Fig. 1.



due to the fact that during cooling the alloy has partially ordered. The two maxima are observed again, the first one being more pronounced. Fig. 1 (c) refers to the alloy in the condition as cooled from 490°C. to 370°C. in 150 hours. The first maximum has now a very high value

indeed, indicating that the ordering process has proceeded to a considerable extent during the annealing period. The Curie point has not changed appreciably.

Above 600°C . the curves a , b , c are the same within the limits of experimental error, suggesting that the specimen returns to the same state (disordered and paramagnetic) and that the effects of previous thermal treatment have been removed.

IV. *Magnetic Measurements.*

The curves in fig. 1 include the effects of both magnetic and order-disorder transformations superimposed on the "normal" curve which would be obtained in the absence of transformations in the range of temperature investigated. Before any quantitative estimate can be made of the thermal effects associated with the order-disorder transformation, it is desirable to determine whether long annealing and its consequent effect in the structure of the alloy is likely to affect the energy involved in the magnetic transformation. Specimens 1.5 mm. in diameter and 4 mm. long were prepared and subjected to heat treatments similar to those given to the specific heat specimen and recorded in the previous section. The saturation intensity was then measured as a function of temperature, the specimens being heated at a uniform rate of 2°C . per minute. These experiments were carried out by Dr. Sucksmith⁽⁷⁾ of Bristol University using the special apparatus he has developed for this type of problem. The results (fig. 2) show that the curve relating saturation intensity and temperature is not greatly affected by annealing. Consequently it is reasonable to conclude * that the energy of the magnetic transformation is to a first approximation independent of the degree of order in the material. The maximum slope on the curves (fig. 2) occurs at about 590°C ., in agreement with the conclusion already mentioned in the previous section that the Curie point is to be associated with the second maximum on the curves of fig. 1.

V. *X-ray Measurements.*

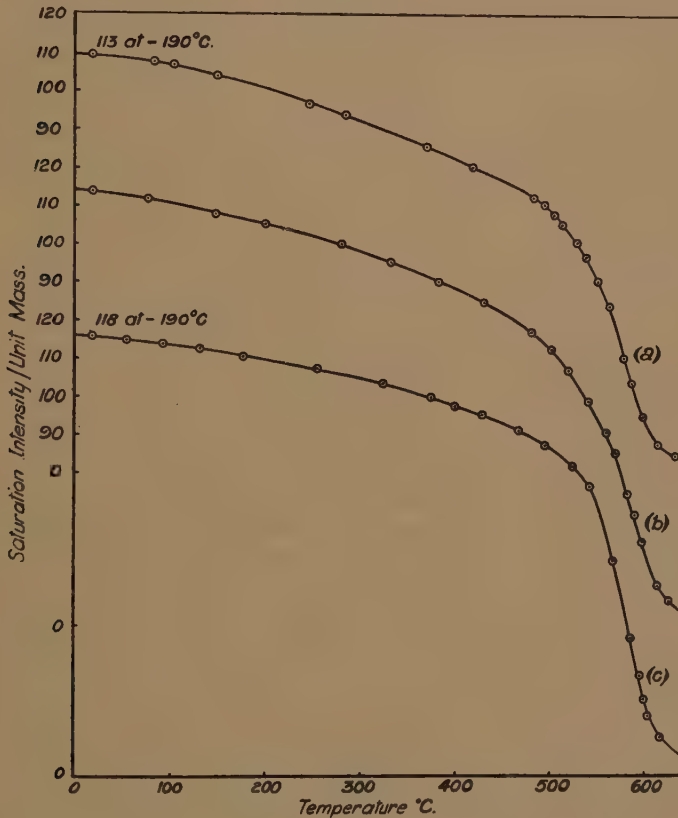
The presence of a superlattice in Ni_3Fe cannot be readily detected due to the similarity in scattering power of iron (26) and nickel (28). The small difference can be enhanced by using a radiation near to the absorption edge of iron, since in this region the scattering power varies in an anomalous manner with wave-length⁽⁸⁾, *e. g.*, for cobalt $\text{K}\alpha$ radiation

* The saturation value of the gram molecular moment of the ordered alloy is about 5 per cent. higher than that of the disordered alloy in the temperature range -190°C . to 100°C . To a first approximation, therefore, the magnetic energy of the ordered alloy will be about 5 per cent. higher than that of the disordered alloy.

the scattering power of iron is considerably depressed (by about 4 units) whilst that of nickel is only slightly affected.

Photographs were taken in a circular camera using the monochromatic $K\alpha$ radiation of cobalt. Films obtained using eight-hour exposures with a specimen of powder as quenched from 700°C . showed no superlattice lines. Specimens annealed for 500 hours in the temperature interval

Fig. 2.



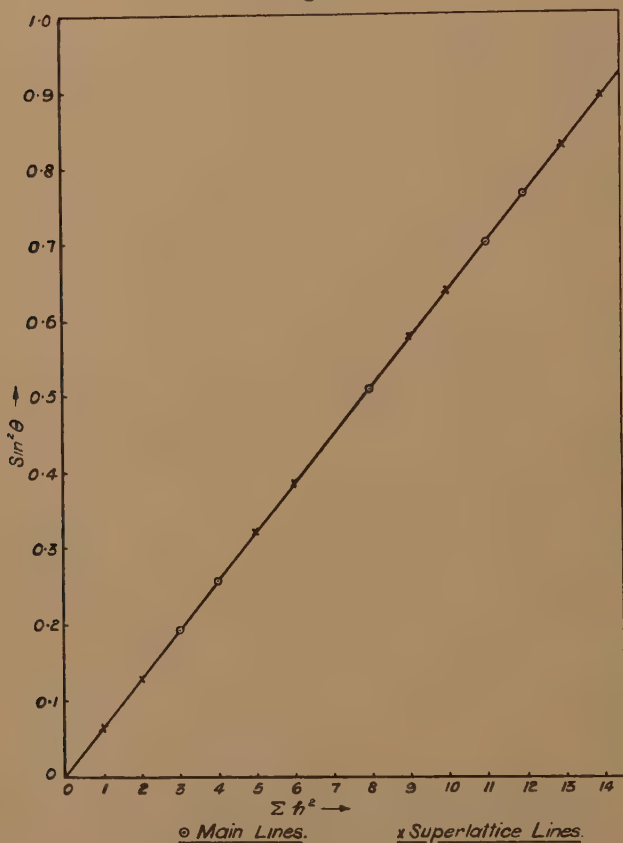
490°C .– 370°C . showed the superlattice lines $\Sigma h^2 = 1, 2, 5, 6, 9, 10, 13, 14$. These were sufficiently strong on films exposed for eight hours for measurements of their position to be made, and the corresponding Bragg reflexion angles (θ) were then calculated. The values of $\sin^2 \theta$ for the various lines are plotted against Σh^2 in fig. 3; all the superlattice lines are present and in the right positions. Both sets of films, on the annealed and quenched alloy, showed two additional lines somewhat fainter than the superlattice lines which are ascribed to impurities introduced during the production of the powder by filing.

The lattice spacings of the annealed and quenched powders were measured using copper $K\alpha$ radiation with an internal filter of nickel. The results obtained were :

as quenched $a = 3.5470 \pm 0.0003$ (3.5472),

as annealed $a = 3.5441 \pm 0.0003$ (3.5464).

Fig. 3.



The difference is considerably greater than the maximum experimental error. The figures in brackets are those given by Bradley ⁽²⁾ for an alloy of similar composition ; the results as quenched are in good agreement ; the difference in the values for the annealed alloy is probably due to the fact that our material was more thoroughly annealed.

Both Kaya and Haworth carried out specific experiments to establish the existence of a superlattice in Ni_3Fe and did not obtain a positive result. In Kaya's case lack of success may be attributed to the fact that the radiation chosen, viz., nickel $K\alpha$, was not particularly suitable.

Haworth chose the $K\beta$ radiation of iron which, according to calculation, should be superior to cobalt $K\alpha$ radiation. Whilst it is doubtful whether any great faith should be placed in the quantitative calculations based on the depression of scattering power in the neighbourhood of the absorption edge, there seems little doubt that the $K\beta$ radiation of iron should have given results in view of the very favourable camera arrangements made by Haworth for the detection of weak lines. Some of his specimens were annealed for 100 hours at $425^\circ\text{C}.$, and although this is not as thorough an anneal as that given to our specimens, it seems probable that it should have given some indication. The nickel content of his specimen was 70 per cent. Experiments carried out on materials of the Ni_3Fe type containing less nickel, or metals other than iron, show that the transformation becomes very much more sluggish as a result of these modifications, and it is conceivable that the specimen used by Haworth would not transform sufficiently in any reasonable period of time for superlattice lines to be observed. (The alloys of the Ni_3Fe type produced commercially exhibit very small thermal anomalies compared with those illustrated in fig. 1 (c) after identical annealing treatments.)

VI. *Comparison with other Order-disorder Transformations.*

Total Energy.—The effect of the order-disorder transformation on the specific heat temperature curve is confined to the temperature interval from $325^\circ\text{C}.$ – $625^\circ\text{C}.$; below $325^\circ\text{C}.$ the rate of diffusion is too low for atomic rearrangement to take place under the experimental conditions used, and above $625^\circ\text{C}.$ the alloy is always disordered. According to the results recorded in section IV. we may assume that the energy of the magnetic transformation is practically independent of the degree of order, and proceed to an estimate of the energy evolved during the order-disorder transformation.

At $325^\circ\text{C}.$, fig. 1 (a), the alloy is disordered, and at $625^\circ\text{C}.$ it has returned again to the disordered state; consequently the energy absorbed in heating the alloy from $325^\circ\text{C}.$ to $625^\circ\text{C}.$ contains no net contribution from the order-disorder transformation, and is equal to that used in increasing the amplitude of the thermal vibrations of the atoms and in changing the alloy from the ferromagnetic to the paramagnetic state. This energy in calories per gram can be obtained from fig. 1 (a) by integration between the temperature limits $325^\circ\text{C}.$ and $625^\circ\text{C}.$ In the case of fig. 1 (c), at $325^\circ\text{C}.$ the alloy is partially ordered, so that a similar integration will give a higher result, as it includes the extra energy required to change the alloy back to the disordered state. The difference in energy content obtained from figs. 1 (a) and 1 (c) is 11.5 cal./gr. and represents the energy given out during annealing as a result of the order-disorder transformation.

The transformation in Ni_3Fe is remarkably sluggish compared with similar transformations in the other alloys which have been examined so far, and it seemed desirable to check whether the treatment given to the specimen, prior to the experiment recorded in fig. 1 (c), was adequate. The annealing treatment was first increased to 200 hours and then to 500 hours, and the energy of the transformation estimated as above was 13.3 cal./gr. and 13.8 cal./gr. respectively. The experimental error is about ± 0.5 cal./gr.; consequently it is reasonable to conclude that the transformation is substantially complete after 500 hours.

Kaya obtained a value of 12.5 cal./gr. on a specimen annealed for 8 days at 490°C . and cooled at 10°C . per day to 450°C . This result is in satisfactory agreement with those recorded above.

According to the Bragg-Williams theory⁽⁹⁾ the total energy E per gram of alloy is given by

$$E = \frac{RT_c}{2.26 M}, \quad \dots \dots \dots (1)$$

where M is the molecular weight of the alloy, T_c the critical temperature in degrees absolute at which the alloy becomes disordered on heating, and R the gas constant. Due to the sluggish character of the transformation, the critical temperature cannot be determined with any degree of certainty from the specific heat temperature curves. Generally this temperature is associated with the temperature at which the specific heat is a maximum; an examination of fig. 1 shows that this temperature increases from 520°C . to 550°C . as the degree of order in the alloy increases. Kaya has shown from resistance measurements that the critical temperature must lie between 490°C . and 520°C ., and for comparison with the Bragg-Williams theory we shall take 520°C ., *i. e.*, 793°K ., as the critical temperature. Substitution in (1) gives the value of E equal to 12.0 cal./gr.

The Bragg-Williams theory is based on the assumption that the energy of a given ordered structure is determined by the average degree of order throughout, *i. e.*, by the long distance or superlattice order. Peierls⁽¹⁰⁾, on the other hand, has carried out a similar calculation based on the assumption that interaction occurred only between nearest neighbours—local order. The total energy released below the critical temperature according to Peierls is about 6 per cent. higher than that given by Bragg and Williams*.

* Haworth, in discussing Kaya's results, takes a value of $E = RT_c/1.33 M$. = 19.0 cal./gr.; this is the result obtained by Peierls for the change from complete order to complete disorder and contains the contribution due to diminution in local order above the critical temperature. Consequently it should not be compared with the experimental values which only refer to energy released up to the critical temperature.

Both results are lower than the experimental value, 15 per cent. and 10 per cent., but the agreement is within the usual limits found in the investigation of other transformations, *e. g.*, Cu_3Au and CuZn . It differs in this respect; for Ni_3Fe the theoretical result is low, whereas for Cu_3Au and CuZn it is high.

The Change in Entropy.—According to Williams ⁽¹¹⁾ the change in entropy accompanying an order-disorder transformation is independent of any assumptions as to the precise mechanism of the ordering process, and should provide a reliable check that the phenomena under consideration are due to the formation of a superlattice. In order to estimate the change in entropy for the Ni_3Fe transformation we have assumed that the dotted curve, fig. 1 (*a*), gives the specific heat temperature curve in the absence of a transformation; it is drawn in such a way that the shaded areas become equal. Using this and the curve obtained on the specimen after 500 hours' annealing, the change in entropy associated with the transformation has been evaluated and found to be 1.14 cal./gram atom. According to the Bragg-Williams theory the value should be $\cdot 56 R$, *i. e.*, 1.12 cal./gr. atom, and somewhat less according to Peierls. The agreement may be considered satisfactory.

Specific Heat above the Critical Temperature.—Due to the variation of local order above the critical temperature, the specific heat is usually higher than the value extrapolated from low temperatures at which no change in atomic configuration takes place. This effect is also present in the specific heat temperature curves of Ni_3Fe , but the additional complication of the magnetic transformation will also introduce a similar effect (the electronic specific heat), and consequently the existence of a high specific heat above the critical temperature cannot be regarded as evidence for an order-disorder transformation.

VII. General.

According to theory the Ni_3Fe transformation should be similar to that in Cu_3Au , and in particular should evolve latent heat at the critical temperature on heating. Although latent heat has not been observed, very high values of specific heat have been recorded (up to 0.7 cal./gr., which is about five times the normal specific heat), and it seems highly probable that much higher values would be obtained if the transformation were not so sluggish.

In Cu_3Au ⁽¹²⁾ it has been shown that the ordering process is retarded very considerably by the formation of antiphase nuclei. There appears to be no reason why this effect should not exist in Ni_3Fe , and our annealing treatments, and apparently Kaya's, were carried out in such a manner as to eliminate this difficulty as much as possible. It was found in the

case of Cu_3Au that, provided the alloy was held just below the critical temperature for a sufficiently long time for the nuclei to grow to a reasonable size (1000 Å.), then the alloy could be cooled fairly rapidly and would still remain in equilibrium. For this reason the Ni_3Fe alloy was always treated at 490° C. for a considerable period (50 hours in the first anneal, fig. 1 (a), and 150 hours in the third anneal) before cooling commenced. The annealing treatment given by Haworth of 100 hours at 425° C. is probably equivalent to less than 10 hours at 490° C.

In experiments on the alloy Cu_3Au as quenched ⁽¹²⁾ the specific heat temperature curve shows two minima, one at 140° C. associated with the formation of a large number of small antiphase nuclei, the other at 320° C. with the growth of these nuclei and the resulting diminution of disordered boundary material. Both processes are governed by diffusion, but the second requires a considerably higher temperature than the first to have appreciable effects.

One minimum only is observed on Ni_3Fe , and it might be argued that the two ordering processes differ in some essential manner. This is not necessarily true. The speed of diffusion in Ni_3Fe will be very much lower than that in Cu_3Au , so that the formation of antiphase nuclei would be expected to occur at a somewhat higher temperature, *i. e.*, the minimum at 425° C. is probably caused by this phase of the ordering process. The growth of nuclei may not become appreciable (under the conditions used for the measurement of specific heat) until the alloy has passed the critical temperature, and, consequently, the effects will not be observed. (The growth of nuclei in Cu_3Au makes no noticeable effect on the specific heat until 240° C. is reached, *i. e.*, 100° C. above the first minimum which occurs at 140° C.)

The specific heat of the Ni_3Fe alloy as quenched is higher * than that of the annealed alloy (4 per cent. \pm 2 per cent.) in the temperature range 200° C.–300° C. It is unusual to find that the normal specific heat of an alloy is dependent on the degree of order; in other alloys the change in specific heat, if any, is less than errors in measurement.

This difference in "normal" specific heat below 300° C. has been neglected in estimating the energy of the order-disorder transformation, since the calculations were limited to the temperature range 325° C.–625° C. As the difference is of the same order as the experimental error, and as it is not known over what range of temperature range it persists, no accurate estimate can be made of the error introduced. It is unlikely to amount to more than 2 cal./gr. and will tend to improve the agreement between the experimental result, 13.8 cal./gr., and the theoretical value of 12 cal./gr.

* This may be attributed to differences in the rate of release of the magnetic energy (*cf.* fig. 2).

VIII. Discussion.

The experimental work of Kaya on specific heat and electrical resistance, together with the work recorded in this paper, indicate that the alloy Ni_3Fe will undergo an order-disorder transformation and form a superlattice. It is not clear how such a transformation, even when suitably controlled, has any specific bearing on the remarkable magnetic properties of the alloy. In order to obtain high permeability in alloys of the Ni_3Fe type it is found necessary in practice to use cooling rates of 1°C . per minute or more through the temperature range 500°C .– 400°C . The binary alloy Ni_3Fe will be by no means completely ordered after this treatment, cf. figs. 1 (b) and 1 (c). In the specimen as cooled at 1°C . per minute, fig. 1 (b), about 30 per cent. of the energy associated with the order-disorder transformation has been released, from which it may be calculated⁽¹²⁾ that the antiphase nuclei, if they exist, will be about 30 Å. in length.

Acknowledgments.

We are indebted to Dr. W. Sucksmith for carrying out the magnetic measurements and to Dr. F. W. Jones for assistance in the X-ray measurements.

The work forms part of the programme of the Structure of Alloys Research Panel of the British Iron and Steel Federation, and was carried out, by kind permission of Dr. A. P. M. Fleming, C.B.E., in the Research Department of the Metropolitan-Vickers Electrical Co., Ltd., Trafford Park.

We are indebted to Professor W. L. Bragg, F.R.S., for his kind interest throughout the investigation.

References.

- (1) Dahl, *Z. f. Metallkunde*, xxviii. p. 133 (1936).
- (2) Bradley, Jay, and Taylor, *Phil. Mag.* xxiii. p. 545 (1937).
- (3) Bradley, *Proc. Roy. Soc. A*, clxvi. p. 353 (1938).
- (4) Kaya, *Journ. of Faculty of Science*, ser. 2, vol. ii. no. 2, p. 31 (1938). Imperial University, Sapporo, Japan.
- (5) Haworth, *Phys. Review*, liv. p. 693 (1938).
- (6) Sykes and Jones, *J. I. M.* lix. p. 257 (1936).
- (7) Sucksmith, *P. R. S.* (in the press).
- (8) Sykes and Jones, *Proc. Roy. Soc. A*, clxi. p. 440 (1937).
- (9) Bragg and Williams *ibid.*, A, cxlv. p. 699 (1934).
- (10) Peierls, *ibid.* A, cliv. p. 207 (1936).
- (11) Williams, *ibid.* A, clii. p. 231 (1935).
- (12) Sykes and Jones, *ibid.* A, clvii. p. 213 (1936).

LXX. *Thermodynamical Properties of some Supraconductors.*

By J. G. DAUNT, A. HORSEMAN, and K. MENDELSSOHN,
Clarendon Laboratory, Oxford.*

[Received January 25, 1939.]

IN 1937 two of us ⁽¹⁾ reported on a determination of equilibrium curves † of some supraconductive substances. From these curves we derived by thermodynamical formulæ ⁽²⁾ the differences of entropy and specific heat between the supraconductive and the normal state which, as we have pointed out, are in most cases very difficult to obtain with sufficient accuracy from direct calorimetric determinations. One of the few metals for which such a determination is possible is tin, and recently Keesom and van Laer ⁽³⁾ have measured its specific heat in the supraconductive and the normal state very accurately. We have thus now the possibility to compare our values, which were derived from purely magnetic measurements, with those obtained in a direct calorimetric determination. The agreement is, as we shall see, excellent, showing that our method is reliable, and we give, therefore, in this paper the equilibrium curves of Indium and Thallium together with the entropy differences and electronic specific heats of these substances. We have also supplemented our determinations on lead and tin by new measurements.

The method of determining the magnetic induction was essentially the same as described previously by one of us ⁽⁴⁾. Below 4° K. the specimens were, however, kept directly in a bath of liquid helium which was carefully stirred. The temperature could be kept constant to less than 0.005° below 1.8° K. and to less than 0.002° between 1.8° and 4° K. The determinations above 4° K. were carried out in the small cryostat described by two of us ⁽⁵⁾. Above 4° K. the temperature could be kept constant within less than 0.01°.

* Communicated by Dr. Kurt Mendelssohn.

† We apply the term "equilibrium" curves to those curves in the H , T diagram which denote the reversible change from zero to normal *induction* of a rod of infinite length in a longitudinal field in contradistinction to the term "threshold" curves which has been reserved for the change of *electrical resistance* under the same conditions. Equilibrium and threshold values coincide sometimes but not always (*cf.* J. G. Daunt, *Phil. Mag.*, in the press). Only the equilibrium values must be applied in thermodynamical considerations.

The Specimens.

The substances employed were, as is necessary in investigations of this kind, of exceptional purity. A thallium sphere was cast *in vacuo* out of "H.S." thallium supplied by Adam Hilger, Ltd. (Laboratory number 7011) of 99.995 per cent. purity. The "frozen in" moment amounted to only 1.4 per cent. The indium specimen was also cast *in vacuo* and had practically spherical shape. The metal was supplied by Johnson, Matthey & Company, Ltd., the purity was given as 99.98 per cent.* The "frozen in" moment was again very small, 0.7 per cent. In the determinations on tin and lead the same specimens of high purity used previously ⁽¹⁾ were employed.

Results.

The equilibrium values are given in Table I. and fig. 1. For thallium and indium the whole curves between 1° K. and the transition point were determined. For lead the part of the equilibrium curve between 4.2° K. and 7.2° K. which was not covered by our previous investigation was measured. In the case of tin only four points at a low temperature were added to our old equilibrium curve ⁽¹⁾ which were needed for a better extrapolation to absolute zero. The temperatures above 4° K. were taken from a gas thermometer and below 4° K. from the vapour pressure of liquid helium †.

Comparing the equilibrium curves with the "threshold" values observed elsewhere we find for thallium that they are in fairly good agreement with the four points measured by de Haas and Voogd ⁽⁸⁾ and for indium with the one point determined by the same authors. The threshold curve for indium determined by Misener ⁽⁹⁾ gives slightly lower values. No values for comparison are available for lead.

The difference in entropy ΔS between the normal (S_n) and the superconductive state (S_s) is given in fig. 2 as a function of temperature. ΔS was calculated as

$$\Delta S = \frac{V}{4\pi} H_c \frac{dH_c}{dT}, \quad (1)$$

* We are indebted to Dr. D. A. Jackson for putting the metal at our disposal.

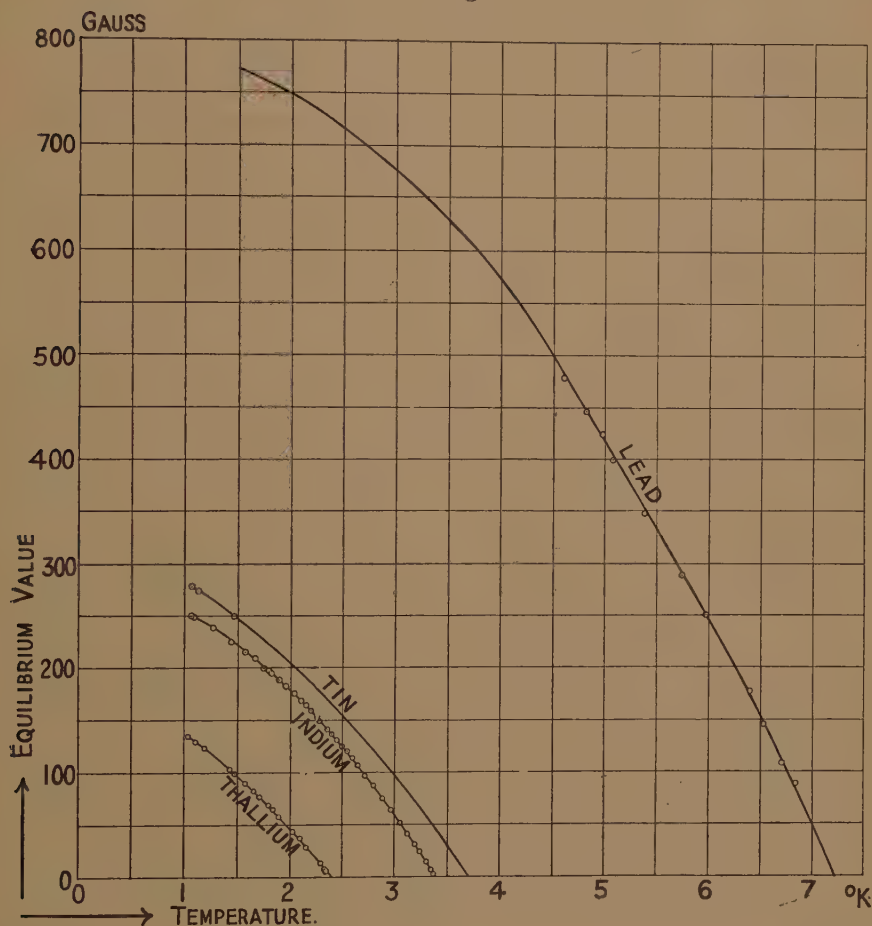
† The temperatures are given in the "1932 scale," ⁽⁶⁾ which is based on a thermodynamic formula. When plotted with the "1937 scale" ⁽⁷⁾ the equilibrium curves showed discontinuities at the λ point. The differentiated curves showed slight kinks near 1.5° K. for both scales, but were somewhat smoother with the "1932 scale." In the ΔS and ΔC curves given in this paper these kinks, which were clearly due to deficiencies in the temperature scales, have been smoothed out. Our results are, however, not appreciably affected by these differences between the 1932 and 1937 scales except, perhaps, for the ΔC curve of thallium.

TABLE I.
Equilibrium values.

T (°K.).	H _c (Gauss).	T (°K.).	H _c (Gauss).
Thallium.		Indium.	
1.04	138.6	2.285	149.5
1.11	133.2	2.357	142
1.20	127	2.404	137
1.24	123.3	2.446	132
1.43	106.4	2.492	126
1.485	101.5	2.538	121
1.58	92.8	2.597	113.7
1.663	85	2.646	107
1.716	79.5	2.72	97.3
1.796	71	2.789	87.8
1.84	66.2	2.886	75.2
1.903	59.25	2.964	63.9
2.03	45.5	3.048	52.3
2.097	38.7	3.122	42.25
2.163	29.5	3.19	32.2
2.282	14.7	3.246	23.4
2.329	8.78	3.31	13.8
2.35	6.12	3.351	7.43
2.361	4.52	3.387	2.66
Indium.		Tin.	
1.07	251	1.065	277
1.1	249.5	1.07	277.5
1.27	239	1.135	274
1.45	225	1.47	250
1.45	226		
1.585	216.5	Lead.	
1.668	210	4.61	478
1.76	201	4.82	447
1.80	198	4.98	424
1.83	196	5.07	399
1.90	189.3	5.38	348
1.965	183.5	5.74	288
2.038	177	5.98	250.5
2.112	169.5	6.40	177
2.156	165	6.53	145.5
2.203	160	6.71	108
2.254	153.5	6.85	88.5

where V is the atomic volume. It can be seen that our determinations of the equilibrium values go to sufficiently low temperatures to allow an extrapolation to absolute zero. We have pointed out ⁽¹⁾ that at sufficiently low temperatures ΔS will be proportional to T and identical

Fig. 1.



with the electronic specific heat in the normal state ($C_n = \gamma T$, according to Sommerfeld's theory). We thus derive γ from the equation

$$\gamma = \left(\frac{d\Delta S}{dT} \right)_{T=0}, \quad \dots \quad (2)$$

The γ values obtained in this way are given in Table II. For tin we arrive at a slightly higher γ value than the one determined previously ⁽¹⁾.

and is given in fig. 3. For the case of tin our values can be compared with the direct determinations by Keesom and van Laer ⁽³⁾. The corre-

Fig. 3.

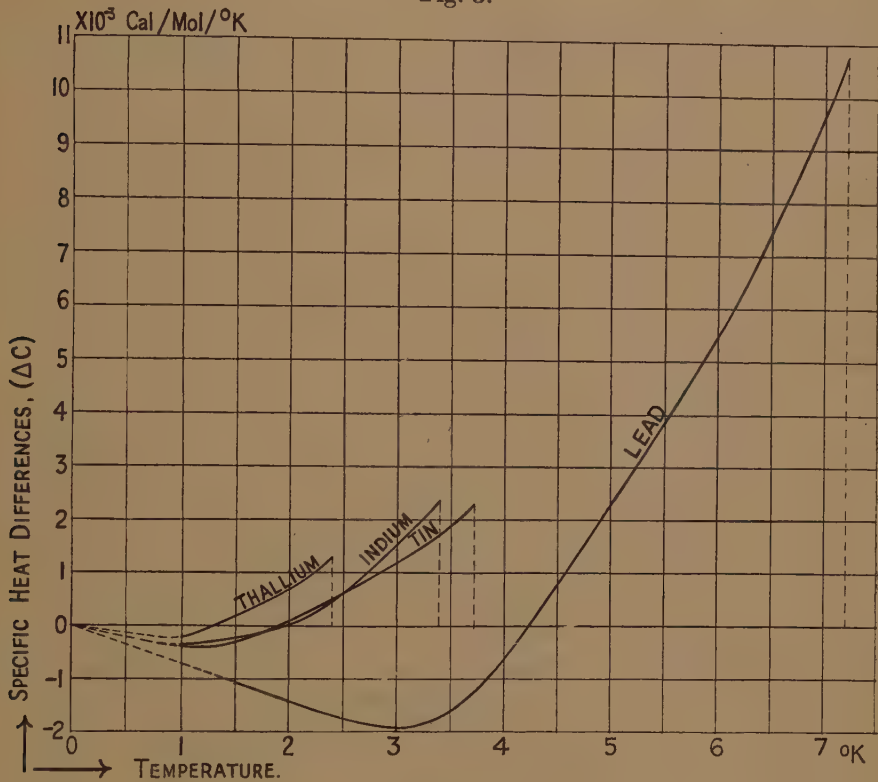


TABLE III.

Specific heat differences in tin.

T.	$\Delta C(10^{-3} \text{ cal./mol/degree}).$	
$^\circ\text{K.}$	Daunt and Mendelssohn (from equilibrium curve).	Keesom and van Laer (direct determination).
1.1 ...	-0.33	-0.35
1.5 ...	-0.22	-0.25
2.0 ...	+0.05	-0.04
2.5 ...	+0.60	+0.60
3.0 ...	+1.17	+1.18
3.3 ...	+1.58	+1.5

sponding values of ΔC given in Table III. show excellent agreement *. The reason is obviously that the determination of the equilibrium curves, being a purely magnetic measurement, can be carried out so exactly that even a double differentiation yields sufficiently accurate results.

The specific heat of the supraconductive electrons ($C_{s.el.}$) can be obtained, as we have shown ⁽¹⁾, from γ and (ΔC) as

$$C_{s.el.} = \Delta C + \gamma T \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Curves showing the dependence of ($C_{s.el.}$) on T are given in fig. 4.

Discussion.

The shape of the ΔS curves at low temperatures leads us to a conclusion on the important question of the number of supraconductive electrons. In attempted theoretical descriptions of the supraconductive state this number has sometimes been assumed to be very small, or of the same order as the number of atoms. Experiments on the threshold value of thin wires carried out by Pontius ⁽¹¹⁾ in this laboratory have made it probable that the latter is the case. This statement is reiterated by a theoretical interpretation of Pontius's results given recently by von Laue ⁽¹²⁾. As it is not probable that the entropy of the supraconductive electrons contains a term that is proportional to $T \uparrow$, the linear change of ΔS with T near absolute zero shows evidently that in this region $\Delta S = S_n$. This means that the entropy of the electrons in the supraconductive state is already practically zero. It is thus evident that at low temperatures the electronic entropy of a supraconductive metal is too small to account for an appreciable number of normal electrons. We can therefore conclude: *in a supraconductive metal at absolute zero all free electrons are in the supraconductive state.*

We have pointed out previously ⁽¹⁾ that the term "free" electrons in this respect not only comprises the valency electrons but all those electrons in the metal which, according to Sommerfeld's theory, have to be treated as "free." For instance, the so-called "positive holes" in transition metals like tantalum and niobium have therefore also to be counted as "free" electrons.

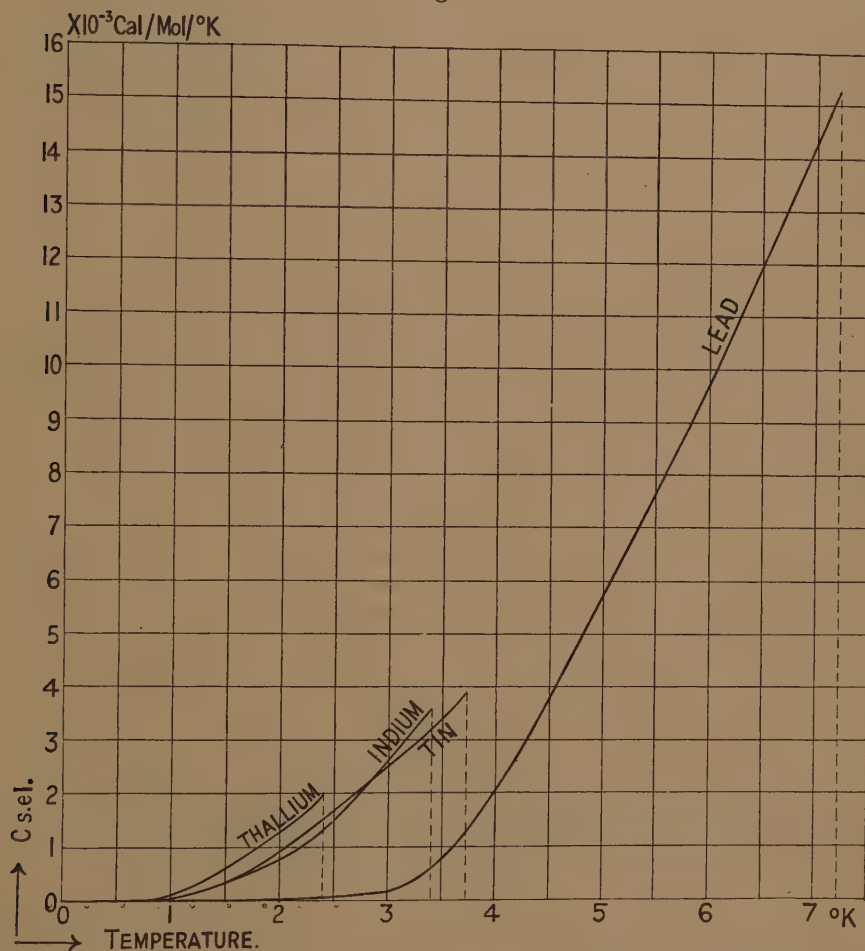
It is evident that the heat taken up by the system of supraconductive electrons in zero field with increasing temperature ($C_{s.el.}$) can be interpreted as a lifting of electrons from the supraconductive into the normal state, thus indicating a gradual change of the proportion of supraconductive to

* The agreement is even better than that given by Keesom and van Laer of their values with the ones from the threshold curve determined by de Haas and Engelkes ⁽¹⁰⁾.

† The experiments by Keesom and van Laer ⁽³⁾ are direct evidence against such a behaviour in the case of tin.

normal electrons from 100 per cent. at absolute zero to 0 at the transition point. Such a conception of a changing "number" of supraconductive electrons has, however, to be treated with great reserve and must not be interpreted as a formation of "electron crystals," as has sometimes

Fig. 4.



been suggested. It needs emphasizing that there exists no phase boundary between the supraconductive and the normal electrons in zero field, and it seems very doubtful whether the lower entropy of the supraconductive electrons in comparison with the "free" electrons can be interpreted as a spatial order. The conception of a variation of the "concentration" of supraelectrons with temperature might prove, however, useful as a

help for a phenomenological treatment as, *e. g.*, the assumption of a change in the number of ferromagnetic electrons with temperature*.

The values for γ which we derive differ somewhat from those calculated by Kok⁽¹³⁾ under the assumption that the number of "free" electrons is equal to the valency. This discrepancy is not surprising, because such an assumption, as has already been pointed out by Sommerfeld⁽¹⁴⁾, cannot be expected to hold, except, perhaps, for the metals in Group I. of the periodic table which are, however, not supraconductive. In the case of the supraconductors a detailed knowledge of the energy spectrum of the metal lattice is necessary for the theoretical evaluation of γ .

By considerations based on Sommerfeld's electronic theory of metals and on F. and H. London's⁽¹⁵⁾ interpretation of supraconductivity, Welker⁽¹⁶⁾ has very recently derived a formula

$$\frac{H_0}{T_c \sqrt{\gamma}} = \text{const.}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

in which H_0 denotes the threshold field at absolute zero and T_c the normal transition point. A connexion between γ , H , and T was pointed out by Kok⁽¹³⁾, and it is also evident from our derivation of γ values from the equilibrium curve that it must be possible to arrive at a relation similar to (5) in a different way. As a first approximation one can assume that all threshold curves are parabolas, hence

$$H = -H_0 \left(\frac{T}{T_c} \right)^2 + H_0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Expressing γ in terms of H and T we arrive at

$$\gamma = \left(\frac{d\Delta S}{dT} \right)_{T=0} = \frac{V}{4\pi} \left(H \frac{d^2 H}{dT^2} \right)_{T=0}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where V is the stomic volume. (6) and (7) give

$$\frac{H_0}{T_c \sqrt{\gamma}} = \sqrt{\frac{2\pi}{V}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

which is very similar to (5). It is interesting that the atomic volume appears in our equation. A combination of (5) and (8)

$$\sqrt{\frac{2\pi}{V}} = \text{const.} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

thus expresses the well-known fact that all supraconductors have similar atomic volume.

In Table IV. numerical values for both sides of equation (8) are given for various supraconductors, the values for H_0 , T_c , and γ having been taken from our equilibrium curves.

* We are indebted to Professor R. Peierls for drawing our attention to this comparison.

SUMMARY.

The equilibrium curves of thallium and indium have been determined between 1° K. and the transition point, and previous determinations on tin and lead have been supplemented by new values.

The differences in entropy and specific heat between the supraconductive and the normal state and the specific heat of the system of supraconductive electrons have been evaluated for all four metals. In the case of tin it could be shown that the values of the specific heat derived from the equilibrium curve agree very well with those obtained by direct calorimetric measurements.

TABLE IV.

	H_0	T_c	$\frac{H_0}{T_c \sqrt{\gamma}}$	$\sqrt{\frac{2\pi}{V}}$
Thallium	170	2.4	.65	.615
Indium	275	3.4	.66	.645
Tin	310	3.7	.62	.625
Mercury	400	4.2	.76	.675
Tantalum	1000	4.4	.80	.75
Lead	800	7.2	.64	.60
Niobium	2590	8.4	.75	.75

It was concluded that at absolute zero *all* free electrons are in the supraconductive state, and that the specific heat of the system of supraconductive electrons might be interpreted as indicating a decrease of the number of supraconductive electrons with temperature.

References.

- (1) J. G. Daunt and K. Mendelssohn, Proc. Roy. Soc. A, clx. p. 127 (1937).
- (2) J. C. Gorter and H. Casimir, *Physica*, i. p. 305 (1934).
- (3) W. H. Keesom and P. H. van Laer, *Physica*, v. p. 193 (1938).
- (4) K. Mendelssohn, Proc. Roy. Soc. A, clv. p. 558 (1936).
- (5) J. G. Daunt and K. Mendelssohn, Proc. Phys. Soc. l. p. 525 (1938).
- (6) W. H. Keesom, Comm. Leiden, Suppl. 71 d.
- (7) G. Schmidt and W. H. Keesom, *Physica*, iv. p. 971 (1937).
- (8) W. J. de Haas and J. Voogd, Comm. Leiden, 212 d.

- (9) A. D. Misener, *Proc. Roy. Soc. A*, clxvi. p. 43 (1938).
- (10) W. J. de Haas and A. D. Engelkes, *Physica*, iv. p. 325 (1937).
- (11) R. B. Pontius, *Phil. Mag.* ser. 7, xxiv. p. 787 (1937).
- (12) M. von Laue, *Ann. Phys.* xxxii. p. 71 (1938).
- (13) J. A. Kok, *Physica*, i. p. 1103 (1934).
- (14) A. Sommerfeld, *Ann. Phys.* xxviii. p. 1 (1937); see also H. Jones and N. F. Mott, *Proc. Roy. Soc. A*, clxii. p. 49 (1937).
- (15) F. and H. London, *Proc. Roy. Soc. A*, cxlix. p. 71 (1935).
- (16) H. Welker, *Ber. Bayr. Akad. Wiss.* p. 155 (1938).

LXXI. *The General Solutions to Problems in Heat Transmission between two Media, one at least of which is a Fluid, the Temperature Gradient for one of the Media along the Line of Flow being known.*

By W. J. WALKER, D.Sc., Ph.D., University of the Witwatersrand, Johannesburg, South Africa *.

[Received February 1, 1939.]

SUMMARY.

The following note is offered as a contribution to the solution of problems in transverse heat transmission, in which the heat transmission is affected by the presence of an extraneous source or sink of large, but unknown capacity. The solution is derived by the assumption of a functional relation between temperature and length of duct for one of the media.

Introduction.

THE problems covered by the somewhat lengthy title of this note are frequently encountered in practice, yet there appears to be no general analysis given in the literature which will enable solutions thereof to be readily obtained. The author has had occasion to attempt solutions in particular cases relating to mine cooling and ventilation †. Since the method followed is applicable in general, the following analysis may be helpful to those who are faced with similar practical problems.

The solutions to the problems relating to the determination of temperatures, along the line of flow, of two fluids flowing either in parallel or counter currents are well known and are covered by the several forms of Grashof's equation. In such cases the solutions are derived by application of the heat balance relation between the two fluid systems. There are in practice, however, many cases where such a heat balance cannot be applied, since, in addition to the transfer of heat from the one medium to the other, there is also an extraneous flow of heat from or to one of the media. The reservoir or sink responsible for this extraneous flow of heat is, in many such cases, of such large capacity as to be responsible for the

* Communicated by the Author.

† "Engineering Problems associated with the Improvement of Temperature and Humidity Conditions of the Atmosphere in Mines at Great Depths." J. H. Dobson and W. J. Walker, *Proceedings Inst. Mech. Eng.*, June (1938), and *Transactions Inst. Chem. Eng.* xvi. Part I. (1938).

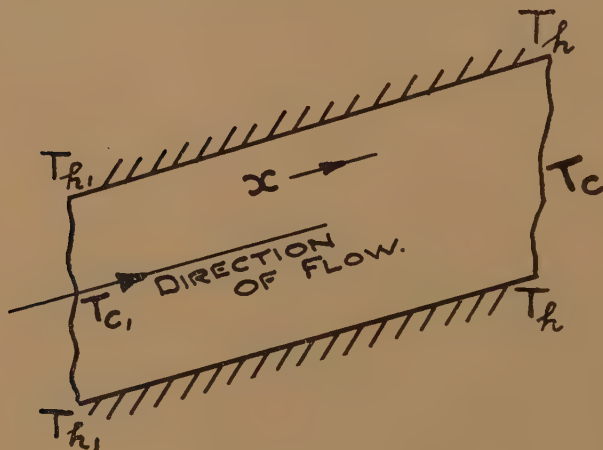
maintenance, during an appreciable period of time, of a practically constant temperature gradient for one of the media along the line of flow. The temperature gradient of the rock along a mine circuit is a case in point, and many others can be visualized.

There are two general cases to be solved, namely, transverse heat flow between a liquid in the one case or a gas in the other, and any other medium the temperature of which along the line of flow is known at all points.

Case 1.—Transverse Heat Transmission between a Liquid and another Medium, the Temperature Gradient of which, along the Line of Flow, is known.

Fig. 1 illustrates the general nature of the problem involved. Here

Fig. 1.



T_{c1} is the initial temperature of the liquid.

T_c is the temperature of the liquid at any point.

T_{h1} is the initial temperature of the other medium.

T_h is the temperature of the other medium at the same cross-section corresponding to T_c .

W_c is the mass of liquid flowing per unit time.

x is the linear co-ordinate along the line of flow. Uniform flow over the cross-section is assumed.

For any elementary peripheral area dA , taken along an elementary length dx , the temperature difference across which in the fluid is dT_c , we have

$$K(T_h - T_c)dA = C_c W_c dT_c \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where K is the heat energy transmitted to the liquid unit area per unit time per unit temperature difference between the liquid and the other medium, and C_c is the specific heat of the liquid.

The suffixes c and h are chosen to refer to the cold and hot media respectively, but these may obviously be interchanged if desired. Since the temperature gradient of the second medium is known along the line of flow, let this be expressed by writing

$$T_h = f(x).$$

If the peripheral dimension of the duct varies, let this be expressed by writing

$$L = \phi(x),$$

where L is the periphery at any cross-section.

Equation (1) then becomes

$$\text{or} \quad K[f(x) - T_c]\phi(x)dx = C_c W_c dT_c$$

$$\frac{dT_c}{dx} + \frac{K \cdot \phi(x)}{C_c W_c} T_c = \frac{K \cdot \phi(x) \cdot f(x)}{C_c W_c},$$

the general solution of which is

$$T_c e^{\frac{K}{C_c W_c} \int \phi(x) dx} = \frac{K}{C_c W_c} \int \phi(x) f(x) e^{\frac{K}{C_c W_c} \int \phi(x) dx} dx. \quad (2)$$

The application of this solution, for example, to the case of a duct system of circular cross-section, and a constant temperature gradient for one of the media in the line of flow, *i. e.*, for

$$\phi(x) = \text{constant} = \pi D,$$

where D is duct diameter

and

$$f(x) = M + Nx = T_h$$

$$\text{gives} \quad T_c = T_h - \frac{W_c C_c N}{\pi DK} - \left[T_{h1} - T_{c1} - \frac{W_c C_c N}{\pi DK} \right] e^{-\frac{\pi DK l}{C_c W_c}}, \quad (3)$$

where l is the length of duct concerned, and T_{h1} and T_{c1} are the initial values of T_h and T_c respectively.

Case 2.—Transverse Heat Transmission between a Gas and another Medium, the Temperature Gradient of which, along the Line of Flow, is known.

Referring again to fig. 1, the differential equation applicable in this case is

$$K(T_h - T_c)dA = C_p W_c dT_c - W_c dx \cos \theta, \quad (4)$$

where C_p is the specific heat of the gas at constant pressure, and θ is the angle which the direction of flow makes with the vertical.

Note that C_p and K are in work or mechanical energy units in this equation.

Using the same symbols as before for T_h and L , and writing $\cos \theta = \phi(x)$, equation (3) becomes

$$K[f(x) - T_c]\phi(x)dx = C_p W_c dT_c - W_c \psi(x)dx,$$

or
$$\frac{dT_c}{dx} + \frac{K \cdot \phi(x)}{C_p W_c} T_c = \frac{K\phi(x)f(x) + W_c \psi(x)}{C_p W_c},$$

the general solution of which is

$$T_c e^{\frac{K}{C_p W_c} \int \phi(x) dx} = \frac{K}{C_p W_c} \int \phi(x) f(x) e^{\frac{K}{C_p W_c} \int \phi(x) dx} . dx + \frac{1}{C_p} \int \psi(x) e^{\frac{K}{C_p W_c} \int \phi(x) dx} . dx. \quad (5)$$

Again, taking the case when

$$\phi(x) = \text{constant} = \pi D,$$

and writing $\theta = 0$ (case of vertical duct) gives

$$T_c = T_h + \frac{W_c}{\pi DK} (1 - C_p N) - \left[T_{h1} - T_{c1} + \frac{W_c}{\pi DK} (1 - C_p N) \right] e^{-\frac{\pi DK h}{C_p W_c}}, \quad (6)$$

where h is the vertical depth of the duct, measured downwards, and T_{h1} and T_{c1} are the initial values of T_h and T_c respectively.

From equations (2), (3), (5) or (6), the value of T_c can be obtained for any point in a system covered by either Case 1 or Case 2. The mean temperature difference along the system may then be calculated in the usual way, thus enabling the heat transmission to be determined.

LXXII. *Errata to Paper "On the Problem of Wave-motion for Sub-Infinite Domains."*—February 1939, p. 182.

To the Editors of the Philosophical Magazine.

GENTLEMEN,—

IN all expressions involving the factor k as a constant multiplier preceding a multiple integral, replace k by the square root of the first factor under the radical sign in the argument of J_1 . This new expression for k is to appear as a factor in the expression of the integrand following the last integral sign. Thus, for instance, the second term of formula 29, page 190, should be rewritten in the form

$$\frac{2x}{\pi^2 a} \int_{\eta=0}^{\infty} \int_{\zeta=-\infty}^{\infty} \iint_{\beta, \gamma=0}^{\infty} \int_{\tau=\frac{x}{a}}^{\infty} \phi_{1,0}(\eta, \zeta, t-\tau) \cdot \sqrt{k^2 - a^2(\beta^2 + \gamma^2)} \\ \cdot \frac{J_1 \left[\sqrt{[k^2 - a^2(\beta^2 + \gamma^2)] \left(\left(\frac{x}{a} \right)^2 - \tau^2 \right)} \right]}{\sqrt{\left(\frac{x}{a} \right)^2 - \tau^2}} \sin \beta y \sin \beta \eta \\ \cdot \cos \gamma(z-\zeta) d\eta d\zeta d\beta d\gamma d\tau.$$

In the statement of the boundary conditions in Tables A, B, and C on pages 182 and 183, the dependence on the variable t has erroneously been left out. Thus, for instance, the functions $\phi(y)$ and $\phi(y, z)$ should be replaced by $\phi(y, t)$ and $\phi(y, z, t)$. Similar remarks apply to Tables B and C.

The term α^2 in the second line of page 186 should be replaced by ξ^2 .

The subscripts 1 in formula 17, on page 188, should be replaced by commas.

The exponential factor $e^{-\frac{(x-\xi)^2}{k}}$, in the last formula of page 194, should be replaced by $e^{-\frac{(x-\xi)^2}{4kt}}$.

Yours faithfully,

ARNOLD N. LOWAN.

May 11, 1939.

LXXIII. *Notices respecting New Books.*

An Introduction to the Theory of Numbers. By G. H. HARDY and E. M. WRIGHT.
(Humphrey Milford, Oxford University Press. Price 25s. net.)

A QUOTATION of two sentences from the authors' preface to this book will indicate a good deal of its special nature. "Our first aim was to write an interesting book." This is the keynote of the work, and the authors must know, as every reader will acclaim, they have succeeded.

The characteristic of the work is that while it maintains accuracy of thought and expression, it is never allowed to become heavy. In many places this causes omissions of extensive theory and result, but the gain in comfort to the reader has made it well worth while, for the theory of numbers has in many parts been uncomfortable of access. This does not mean that good text-books on the theory have been lacking; on the contrary, there has been no lack of them, but in general these books have been content to attract by the grim insistence upon detailed sequence and accuracy, rather than by the ready presentation of the high colours of the theory. That is to say, books have usually been written in a manner oriented specially for those with exceptional powers for research in the field, rather than with the object of throwing open the joys of the subject, which we all know to possess peculiar attractions for a wide range of persons outside the relatively small group of successful research workers in the field itself. This book will provide a view of the subjects treated therein for many a young student; for those whose abilities falter before more formal treatment this book will open new doors, and for all readers its beauty of treatment will be a source of pleasure.

As the book will facilitate the initial approach to many subjects in the theory of numbers, it is accordingly difficult to select subjects for individual mention in this review; doubtless such selections from the book would vary considerably with the selector.

There is a block of three chapters dealing with congruence theory, the first two of which form a very readable and well exemplified introduction to this subject-matter. Fermat's theorem and Quadratic Reciprocity are attractively presented. The account of continued fractions is an excellent one, and it is a contribution to mathematical literature in English; it is comprehensive, easy to read, and has a delightful elegance. The chapters on arithmetical functions, their generating functions and orders of magnitude make excellent reading. There are four very good chapters on algebraic number theory which introduce that subject by means of particular matters—a very excellent mode of introduction in the opinion of the reviewer. Most of these chapters are likely to be within the profitable approach of successful open scholarship candidates, and one likes to think of the brilliance of the subject-matter unfolding before young eyes.

There are sections of the book where the authors have been forced to omit much heavy analysis in their resolve to be interesting. This is true in work upon Waring's and similar problems. The authors have set out known results

and have given a structural sketch of proofs. One can easily imagine experts in other mathematical fields reading these sections with enjoyment, for the results themselves are frequently only hazily known, and it is interesting to know the structural line of approach, even though the ingenuities for solving the enormous difficulties of modern analysis are still for the specialized expert only.

The printing and setting out of the book are in keeping with its essential attractiveness of literary construction, and is fully up to our expectation from printers and publishers. It forms a most attractive volume.

A Dual Theory of Conduction in Metals. By EDWIN HERBERT HALL.
[Pp. viii+170.] (Cambridge, Mass.: The Murray Printing Company, 1938. Price \$2.00.)

As one of his last contributions to a field which he so enriched by his own investigations, the discoverer of the Hall effect has here presented a connected account of ideas tentatively developed over a period of some twenty years. The general background is that of the classical theory of conduction. "I do not criticize those who choose to attack the problems of metallic conduction by the methods of wave mechanics," Hall says, "nor do I belittle the results they may have attained. I merely present a summary of my own ideas and my own reasoning to be appraised by the infallible judgment of time." Although some readers may be excusably irritated by the apparent indifference to recent theoretical methods, in a field so complex any effective coordination is of potential value. It is with coordination that this book is primarily concerned, and many of the relations found are to a large extent independent of the particular mechanical model which has led to their discovery.

According to Hall's dual theory, electrical conduction is maintained in part by free electrons, and in part by "associated electrons . . . which pass from atomic union to atomic union . . . without acquiring energy of thermal agitation." The mathematical postulates of the treatment are embodied in three equations, expressing the form of the relation between the number of free electrons and the temperature, the temperature dependence of the ratio of the free electron to the total conductivity, and the temperature dependence of the heat of ionization. These equations give seven characteristic constants for each metal which are determined from empirical data, and then used in making predictions which may be compared with observation.

In the light of the general theory, discussions are given of electrical and thermal conductivity, the various thermo-electric effects, photoelectric and thermionic emission, the Volta effect, and, finally, the magnetic transverse effects. The experimental data are critically reviewed, and there are many useful tables. In appendices short accounts are given of the author's views on superconductivity, optical absorption in metals, cold emission, and metals under pressure.

The author would have been far from claiming that his treatment gives a satisfactory explanation of the various phenomena. Indeed, the most prominent aim in the book is at empirical coordination rather than at explanation. For its ordering and presentation of extensive experimental material from a point of view reached by long and intensive investigation of the problem of conductivity, this book is well worthy of study by all those concerned with the electrical properties of metals.

E. C. S.

Superconductivity. By D. SHOENBERG. [Pp. x+112.] (Cambridge University Press, 1938. Price 6s.)

IN spite of the wide successes of the electronic theory of the metallic state, the phenomena of superconductivity still seem to fall outside its range. The absence of an even tentative theoretical point of view makes the ordering of the experimental facts much more difficult. None the less, the advances of recent years have been such as to enable the author to succeed remarkably well in his aim at giving "a coherent statement of what it is that the theory has to explain." Foremost among the advances are those initiated by the investigations of Meissner and Ochsenfeld, which showed that a superconductor is characterized fundamentally rather by zero magnetic induction than by zero resistivity. Further, the application of thermodynamics to the treatment of the transition from the normal to the superconducting state as a phase change has resulted in an extensive coordination, and a wide range of properties is effectively covered by a specification of the temperature variation of the critical field. The ordered description of superconductivity is therefore appropriately centred on the account of the magnetic properties of superconductors, and on the treatment of the thermodynamics of the transition process.

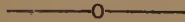
After an outline of the essential facts, the magnetic properties of a superconductor are considered, being contrasted with those of a perfect conductor. As with ferromagnetics, shape effects are important, and since the field will not, in general, be uniform throughout a specimen, it is necessary to consider the intermediate state which occurs when the critical field is exceeded over only part of a specimen. The disturbance of superconductivity by a current and the persistent current phenomena in multiply connected conductors are discussed. The thermodynamics of superconductivity is adequately surveyed, and an account is given of superconducting alloys, and of investigations on thin films, the further study of which is likely to be very profitable.

An appendix provides an up-to-date and well-documented summary of experimental data, giving transition temperatures and critical field-temperature curves for elements, and data for superconducting compounds and alloys.

This clear survey serves admirably the general purpose of the series in providing an authoritative account of the present position in a subject of topical physical interest. It may be recommended to all those who are interested in an unsolved and clearly fundamental problem of modern physics. E. C. S.

[The Editors do not hold themselves responsible for the views expressed by their correspondents.]

INDEX TO VOL. XXVII.



APPLETON (E. V.) and Naismith (R.), the variation of solar ultra-violet radiation during the sun-spot cycle, 144.

Approximation formulæ for a well-known difference of products of two cylinder functions, 407.

Atomic heat of potassium, 551.

Auluck (F. C.), entropy of Fermi-Dirac gas, 370.

Auxiliary equation for use with the Heaviside expansion theorem, 404.

Banerjee (S. S.) on the critical dimensions of tuned transmitting circular loop aerials, 174.

Basu (D.) and Kar (K. C.) on the neutron-proton scattering, 76.

Bell (D. A.) on the general validity of Nyquist's theorem, 645.

Benzene, dielectric constant of, 661.

Bikerman (J. J.) and Rideal (E. K.), a note on the nature of sliding friction, 687.

Biot (M. A.), non-linear theory of elasticity and the linearized case for a body under initial stress, 468.

Blumenthal (E.), theory of the fractionation of gaseous mixtures by diffusion. The characteristics of the Hertz-mercury-vapour-pump, 341.

Books, new :—M. M. Fréchet, *Traité du Calcul des Probabilités et de ses Applications*. Tome i. fascicule 3: *Recherches Théoriques Modernes*, 129; W. Edwards Deming, *Some Notes on Least Squares*, 130; A. J. Ansley, *An Introduction to Laboratory Tech-*

nique, 130; C. N. Moore, *Summable Series and Convergence Factors*, 131; D. G. Fink, *Engineering Electronics*, 131; P. Wessel, *Physik für Studieren an Technischen Hochschulen und Universitäten*, 132; E. Hirschclaff, *Fluorescence and Phosphorescence*, 132; F. E. Hoare, *A Textbook of Thermodynamics*, 258; W. M. Smart, *Stellar Dynamics*, 258; E. Sutton, *Demonstration Experiments in Physics*, 260; V. M. Faires, *Applied Thermodynamics*, 386; V. M. Faires, *Elementary Thermodynamics*, 386; D. Jordan Lloyd and D. Shore, *Chemistry of Proteins*, 387; L. Bergmann, *Ultrasonics and their Scientific and Technical Applications*, 513; R. C. Tolman, *The Principles of Statistical Mechanics*, 514; C. E. van Horn, *A Preface to Mathematics*, 515; J. M. Thomas, *Theory of Equations*, 515; A. A. Albert, *Modern Higher Algebra*, 515; F. Daniels, *Lectures on Chemical Kinetics*, 516; A. Findlay, *The Phase Rule and its Applications*, 644; J. C. Speakman, *Modern Atomic Theory*, 644; G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 770; E. H. Hall, *A Dual Theory of Conduction in Metals*, 771; D. Shoenberg, *Superconductivity*, 772.

Bosson (G.), the flexure of an infinite elastic strip on an elastic foundation, 37.

- Bremmer (H.) and Pol (Balth. van der), further note on the propagation of radio waves over a finitely conducting spherical earth, 261.
- Buchholz (H.), approximation formulæ for a well-known difference of products of two cylinder functions, 407.
- Burington (R. S.) and Risley (N. S.), a matrix theory development of the theory of symmetric components, 605.
- Calculation of equilibrium internuclear distance for diatomic hydrogen, hydrides, and deuterides in ground and excited states, 389.
- of triode constants, 709.
- Carpenter (L. G.) and Steward (C. J.), the atomic heat of potassium, 551.
- Cattermole (J.) and Wilson (W.), the elementary particle, 84.
- Clark (C. H. D.) and Stoves (J. L.), the calculation of equilibrium internuclear distance for diatomic hydrogen, hydrides, and deuterides in ground and excited states, 389.
- Colwell (R. C.), Stewart (J. K.), and Friend (A. W.), symmetrical figures on circular plates and membranes, 123.
- Concept and determination of mass in Newtonian mechanics, 33.
- Coulthard (W. B.), an auxiliary equation for use with the Heaviside expansion theorem, 404.
- Critical dimensions of tuned transmitting circular loop aërials, on the, 174.
- Dasannacharya (B.) and Seth (Amar Chand), Geiger point counters, 249.
- Daunt (J. G.), Horseman (A.), and Mendelssohn (K.), thermodynamical properties of some superconductors, 754.
- Davy (N.), Littlewood (J. H.), and McCaig (M.), the force-time law governing the impact of a hammer on a stretched string, 133.
- Development of turbulent liquid motion over an infinite plate, on the, 240.
- Dielectric constant, dipole moment, and molecular polarization of 1·4 dioxane ($C_4H_8O_2$), 669.
- — of benzene, 661.
- Dielectrics, thermal instability of, for alternating voltages when the loss angle is dependent upon the field strength, 582.
- Diffraction and refraction of a horizontally polarized electromagnetic wave over a spherical earth, 421.
- of wireless waves round the earth, 517.
- Dingle, (H.), the interpretation of the Michelson-Morley and Kennedy-Thorndike experiments, 693.
- Dioxane, 1·4 ($C_4H_8O_2$), dielectric constant, dipole moment, and molecular polarization of, 667.
- Direct determination of stresses from the stress equations in some two-dimensional problems of electricity. — Part II. Thermal stresses, 437. Part III. Problems of non-isotropic material, 596.
- Disintegration of erythrocytes and denaturation of hemoglobin by high pressure, 637.
- Dixon's formula for well-poised series, on, 579.
- Dow (R. B.) and Matthews, Jr. (J. E.), the disintegration of erythrocytes and denaturation of hemoglobin by high pressure, 637.
- Elasticity, non-linear theory of, and the linearized case of a body under initial stress, 468.
- Electronic waves, 1.
- Elementary particle, the, 84.
- Entropy of Fermi-Dirac gas, 370.
- Equilibrium diagram of iron-nickel alloys, 614.
- Equivalence of two piezoelectric oscillating quartz crystals of symmetrical outlines with respect to a plane perpendicular to an electrical axis, 640.

- Errata to paper "On the problem of wave-motion for sub-infinite domains," 769.
- Euler's problem of two fixed centres of gravitation, 149.
- Evidence for a superlattice in the nickel-iron alloy Ni_3Fe , 742.
- Factors affecting the limit of solubility of elements in copper and silver, 294.
- Finite strain, some problems of.—I., 286; II., 449.
- Fischer (Otto F.), Hamilton's quaternions and Minkowski's potentials, 375.
- Flexure of an infinite elastic strip on an elastic foundation, 37.
- Force-time law governing the impact of a hammer on a stretched string, 133.
- Formulae for the associated Legendre functions of the first kind, 703.
- Fremlin (J. H.), calculation of triode constants, 709.
- Friction, sliding, a note on the nature of, 687.
- Friend (A. W.), Colwell (R. C.), and Stewart (J. K.), symmetrical figures on circular plates and membranes, 123.
- Geiger point counters, 249.
- Gemant (A.) and Whitehead (S.), thermal instability of dielectrics for alternating voltages when the loss angle is dependent upon the field strength, 582.
- General solutions to problems in heat transmission between two media, one at least of which is a fluid, the temperature gradient for one of the media along the line of flow being known, 765.
- Geometric derivation of the second order wave equation, 221.
- Geometry of discrete vector maps, 98.
- Geophysical consolidation problems, on some, 576.
- Gilbert (C.), a kinematical description of a flat space-time, 543.
- Graphical construction of stress, 445.
- Gravitation, quantum theory of, 453.
- Gray (M. C.), diffraction and refraction of a horizontally polarized electromagnetic wave over a spherical earth, 421.
- Griffith (E.) and Sherratt (G. G.), a hot wire method for the thermal conductivities of gases, 68.
- Guest (J. J.), a graphical construction for stress, 445.
- Hamilton's quaternions and Minkowski's potentials, 375.
- Haskey (H. W.), a geometric derivation of the second order wave equation, 221.
- Heat conduction, note on a problem in, 706.
- Heaviside expansion theorem, an auxiliary equation for use with the, 404.
- Hertz-mercury-vapour-pump, the characteristics of the, 341.
- Horseman (A.), Daunt (J. G.), and Mendelssohn (K.), thermodynamic properties of some superconductors, 754.
- Hot wire method for the thermal conductivities of gases, 68.
- Hydrodynamic forces on an accelerated cylinder moving in two-dimensions, 200.
- Interpretation of the Michelson-Morley and Kennedy-Thorndike experiments, 693.
- Iron-nickel alloys, equilibrium diagram of, 614.
- Jones (E. Taylor), optical dispersion and the vibratory doublet photon, 565.
- Kaliviaris (A.), quantum theory of gravitation, 453.
- Kar (K. C.) and Basu (D.) on the neutron-proton scattering, 76.
- Kinematical description of a flat space-time, 543.
- Koga (Issac), equivalence of two piezoelectric oscillating quartz crystals of symmetrical outlines with respect to a plane perpendicular to an electrical axis, 640.

- Kothari (D. S.), three elementary examples of the uncertainty principle, 62.
- Leech (P.) and Sykes (C.) the evidence for a superlattice in the nickel-iron alloy Ni_3Fe , 742.
- Legendre functions of the first kind, some formulæ for the associated, 703.
- Lift and drag of a rectangular wing spanning of a free jet of circular section, 229.
- Littlewood (J. H.), Davy (N.), and McCaig (M.), the force-time law governing the impact of a hammer on a stretched string, 133.
- Lowan (A. N.) on the problem of wave-motion for sub-infinite domains, 182.
- , errata to paper "On the problem of wave-motion for sub-infinite domains," 769.
- Löwy (H.) on some geophysical consolidation problems, 576.
- McCaig (M.), Davy (N.), and Littlewood (J. H.), the force-time law governing the impact of a hammer on a stretched string, 133.
- MacRobert (T. M.) on Dixon's formula for well-poised series, 579.
- , some formulæ for the associated Legendre functions of the first kind, 703.
- Magnesium alloys, the thermal and electrical conductivities of some, 677.
- Martin (M. H.), Euler's problem of two fixed centres of gravitation, 149.
- Mass in Newtonian mechanics, the concept and determination of, 33.
- — —, a further note on the definition and determination of, 51.
- Mathur (Sukhdeo Bihari), a note on the method of parallax and the resolving power of the eye, 94.
- Matric theory development of the theory of symmetric components, 605.
- Matthews, Jr. (J. E.) and Dow (R. B.), the disintegration of erythrocytes and denaturation of hemoglobin by high pressure, 637.
- Mendelssohn (K.), Daunt (J. G.), and Horseman (A.), thermodynamical properties of some superconductors, 754.
- Millington (G.), the diffraction of wireless waves round the earth (a summary of the diffraction analysis, with a comparison between the various methods), 517.
- Mokhtar (M.), studies of the tone quality of organ pipes.—II. Reed-pipes, 195.
- Naismith (R.), and Appleton (E. V.), the variation of solar ultra-violet radiation during the sunspot cycle, 144.
- Narlikar (V. V.), the concept and determination of mass in Newtonian mechanics, 33.
- Neutron-proton scattering, on the, 76.
- Non-linear theory of elasticity and the linearized case for a body under initial stress, 468.
- Note on a problem in heat conduction, 706.
- — — the method of parallax and the resolving power of the eye, 94.
- — — the nature of sliding friction, 687.
- Nyquist's theorem, on the general validity of, 645.
- Obituary.—Professor A. W. Porter, 388.
- Ôkubo (H.), the stress distribution in an aeolotropic circular disk, 508.
- Omara (M. A.), hydrodynamic forces on an accelerated cylinder moving in two-dimensions, 200.
- Optical dispersion and the vibratory doublet photon, 565.

- Owen (E. A.) and Roberts (the late E. Wynne), factors affecting the limit of solubility of elements in copper and silver, 294.
- and Sully (A. H.), the equilibrium diagram of iron-nickel alloys, 614.
- Parallax and the resolving power of the eye, a note on the method of, 94.
- Pendse (C. G.), a further note on the definition and determination of mass in Newtonian mechanics, 51.
- Pol (Balth. van der) and Bremmer (H.), further note on the propagation of radio waves over a finitely conducting spherical earth, 261.
- Powell (R. W.), the thermal and electrical conductivities of some magnesium alloys, 677.
- Propagation of radio waves over a finitely conducting spherical earth, further note on the, 261.
- Quantum theory of gravitation, 453.
- Ray (Manohar) on the development of turbulent liquid motion over an infinite plate, 240.
- Rideal (E. K.) and Bikerman (J. J.), a note on the nature of sliding friction, 687.
- Risley (N. S.) and Burington (R. S.), a matrix theory development of the theory of symmetric components, 605.
- Roberts (the late E. Wynne) and Owen (E. A.), factors affecting the limit of solubility of elements in copper and silver, 294.
- Sen (Bibhutibhusan), direct determination of stresses from the stress equations in some two dimensional problems of elasticity. —Part II. Thermal stress, 437; Part III. Problems of non-isotropic material, 596.
- Seth (Amar Chand) and Dasanna-charya (B.), Geiger point counters, 249.
- Seth (B. R.), some problems of finite strain.—I., 286; II., 449.
- , uniform motion of a sphere or a cylinder through a viscous liquid, 212.
- Sherratt (G. G.) and Griffiths (E.), a hot wire method for the thermal conductivities of gases, 68.
- Solar ultra-violet radiation during the sunspot cycle, variation of, 144.
- Squire (H. B.), the lift and drag of a rectangular wing spanning of a free jet of circular section, 229.
- Stewart (C. J.) and Carpenter (L. G.), the atomic heat of potassium, 551.
- Stewart (J. K.), Colwell (R. C.), and Friend (A. W.), symmetrical figures on circular plates and membranes, 123.
- Stoves (J. L.) and Clark (C. H. D.), the calculation of equilibrium internuclear distance for diatomic hydrogen, hydrides, and deuterides in growing and excited states, 389.
- Stress, graphical construction for, 445.
- distribution in an aeolotropic circular disk, 508.
- Stresses from the stress equations in some two-dimensional problems of elasticity, direct determination of.—Part II. Thermal stresses, 437; Part III. Problems of non-isotropic material, 596.
- Sully (A. H.) and Owen (E. A.), the equilibrium diagram of iron-nickel alloys, 614.
- Superlattice in the nickel-iron alloy Ni_3Fe , the evidence for, 742.
- Sykes (C.) and Leech (P.), the evidence for a superlattice in the nickel-iron alloy Ni_3Fe , 742.
- Symmetrical components, matrix theory development of the theory of, 605.
- figures on circular plates and membranes, 123.

- Theorem of three moments with variable "I" but without thrust, 328.
- Theory of the fractionation of gaseous mixtures by diffusion. The characteristics of the Hertz-mercure-vapour-pump, 341.
- Thermal and electrical conductivities of some magnesium alloys, 677.
- conductivities of gas, hot wire method for the, 68.
- instability of dielectrics for alternating voltages when the loss angle is dependent upon the field strength, 582.
- stability of a cylindrical stratified dielectric, 276.
- Thermodynamical properties of some supraconductors, 754.
- Thomson (J. J.), electronic waves, 1.
- Three moments with variable "I" but without thrust, theorem of, 328.
- Tone quality of organ pipes, studies of the.—II. Reed-pipes, 195.
- Triode constants, calculation of, 709.
- Turton (F. J.), the theorem of three moments with variable "I" but without thrust, 328.
- Uncertainty principle, three elementary examples of the, 62.
- Uniform motion of a sphere or a cylinder through a viscous liquid, 212.
- Variation of solar ultra-violet radiation during the sunspot cycle, 144.
- Vaughan (W. C.), the dielectric constant of benzene, 661.
- , dipole moment, and molecular polarization of 1.4 dioxane ($C_4H_8O_2$), 669.
- Vector maps of finite and periodic point sets, 490.
- Walker (W. J.), the general solutions to problems in heat transmission between two media, one at least of which is a fluid, the temperature gradient for one of the media along the line of flow being known, 765.
- Wave-motion for sub-infinite domains, on the problem of, 182.
- Whitehead (S.), the thermal stability of a cylindrical stratified dielectric, 276.
- and Gemant (W.), thermal instability of dielectrics for alternating voltages when the loss angle is dependent upon the field strength, 582.
- Wilson (W.) and Cattermole (J.), the elementary particle, 84.
- Wireless waves round the earth, diffraction of, 517.
- Woolley (H. W.), note on a problem in heat conduction, 706.
- Wrinch (D. M.), the geometry of discrete vector maps, 98.
- , vector maps of finite and periodic point sets, 490.

END OF THE TWENTY-SEVENTH VOLUME.